Path Invariants and the Evolution of Replication

by

Brian K. Davis

Research Foundation of Southern California

P.O. Box 13595, La Jolla, CA 92039, U.S.A.

(*published December 31, 2015*)

Summary. Carbohydrate participation in nascent nucleotide synthesis pathways and transition states in ribozyme catalyzed reactions were proposed to represent the conserved vestiges of a pre-RNA metabolism. Invariant features of ketopentols and related metabolites, principally in the CO_2 driven autocatalytic reductive pentose-phosphate cycle, furthermore, indicate pre-RNA metabolism was centered on increasingly advanced polycarbotide replicators. They broadly spanned an interval from spontaneous glycoaldehyde autocatalysis to RNA. Linear polycarbotide sequences could encode algorithms for structures and functions promoting self-propagation and accumulate complexity. Additionally, they served to stabilize labile metabolites and acted as a chiral filter; identifying a source for *in vivo* homochirality of sugars and, plausibly, amino acids. A tandem distribution of ribotide- and ribotide-acylating carboxyl invariants among pyrimidine synthesis intermediates indicated an exchange of polycarbotide cofactors, conducive to replication with nucleobase intermediates, during formation of the reaction sequence. Retention of a ribose-5'-phosphate platform by intermediates in purine synthesis similarly conformed with initial attachment of a polycarbotide cofactor. Consistent with nucleobase precursor amino acid participation in pre-RNA metabolism, they acquired codons from ancient triplet clusters within the genetic code. Surface-bound, self-annealed polycarbotides were also equipped to circumvent the de Duve-Miller restrictions on a primal surface metabolism. Model carbozymes were constructed of the Tamura-Schimmel synthetase and early peptidyl transferase A- and P-site ribozymes. Replication was revealed to have evolved by accretion, where each successive replicator built upon its predecessor, in a series of 'piggy-back'-like transitions. The proposed model of pre-RNA evolution also led to a mechanism coupling geophysical free energy with the self-organizing capacity of polycarbotide replication.

Key words: pre-RNA metabolism, polycarbotide replication; chiral filter; encoded algorithms; de Duve-Miller restrictions; 'piggy-back' transitions; coupled free energy.

Abbreviations

A	adenine
ACS	acetyl-coenzyme A synthase
Ala	alanine
Asn	asparagine
Asp	aspartate
bp	base pair
C	cytodine
CABP	2-carboxy-arabinose-1,5-bisphosphate
CKABP	2-carboxy-3-keto-arabinose-1,5-bisphosphate
CoA	coenzyme A
CODH	carbon monoxide dehydrogenase
Co(I/III)FeSP	cobalt (Co^{1+}, Co^{3+}) bearing corrinoid iron sulfur protein
CT	central trunk
Cys	cysteine
DCHB	dicarboxylate/4-hydroxybutyrate cycle
DHAP	dihydroxy-acetone-3-phosphate
EP	erythrose-4-phosphate
Fd	ferredoxin
G	guanine
Glu	glutamate
Gln	glutamine
Gly	glycine
GAP	glyceraldehyde-3-phosphate
GP	glycerol-3-phosphate
HPBC	3-hydroxypropionate bi-cycle
HPHB	hydroxypropionate/4-hydroxybutyrate cycle
Ile	isoleucine
KG	α-ketoglutarate
Leu	leucine
N	standard nucleobase
nt	nucleotide
Orn	ornithine
OA	oxaloacetate
PP	pyrophosphate
Pro	proline
PGA	3-phospho-glycerate
Pyr	pyruvate
R	purine
RACA	reductive acetyl-coenzyme A
RCC	reductive citrate cycle
RDP	ribose-5-diphosphate
rf	replicative form
RP	ribose-5-phosphate
RPC	reductive pentose-phosphate cycle
RTP	ribose-5-triphosphate
RuBP	ribulose-1,5-bisphosphate
RuP	ribulose-5-phosphate
Ser	serine
THF	tetrahydrofolate
THMP	tetrahydrometanopterin
Thr	threonine
Val	valine
U	uridine
UQ	ubiquinone
X	coding-site nucleobase
Y	pyrimidine

Notation: amino acid superscripts, synthesis path-length; square brackets, a monomer or scaffold invariant; circle with arrow, single rotation of cyclical pathway

Terms: carbotide, a sugar-phosphate; ribotide, ribose-phosphate; carbozyme, catalytic carbotide; polypentotide, RPC polymeric form; synthesome, polymer complex for amino acid synthesis; replisome, polymer complex for replication.

1. Introduction

Ribulose-1,5-bisphosphate (RuBP) molecules linked by parallel polyphosphate scaffolds were recently implicated in the pre-RNA replication of polymeric 3-phosphoglycerate [PGA] (Davis 2012). Motivating this proposition were findings from a series of investigations undertaken to clarify the origin of the genetic code, particularly with respect to the role of 'pre-species-divergence' tRNA in its formation (Davis, 1999, 2006, 2008, 2009, 2011). They linked α-carboxyl ubiquity among intermediates in amino acid synthesis pathways to the attachment of a bifunctional tRNA cofactor/ adaptor molecule during code formation, in the transition from the RNA era of life (Orgel, 1968; White, 1972; Gilbert, 1986). Adding credence to the concept that path invariants were former cofactor attachment sites, these investigations further revealed that conspicuous anomalies in codon assignments, within the code, resulted from an exchange of tRNA cofactors during synthesis, based on the distribution of dicarboxylated amino acid intermediates and sequence identity between pre-divergence tRNA species (Davis, 2013a). An invariant phosphate at each end of components in the reductive pentose cycle (RPC), by extension, provided evidence of early attachment to dual polyphosphate scaffolds. 3-Phospho-glycerate [PGA] synthesis, in this context, implied a simple (polycarbotide) replicator preceded RNA, as anticipated by Orgel (2008), and that replication coevolved with the pathways of central metabolism. Distinct from a number of stereochemically motivated model replicators, including ANA (Diederichsen, 1996), GNA (Zhang et al., 2005), homo-DNA (Groebke et al., 1988), pRNA (Pitsch et al., 2003), PNA (Egholm et al., 1992), and TNA (Schoning et al., 2000), the polycarbotide replicator attributed to the RPC is a simple antecedent of RNA, deduced from structural invariants in this C-fixing pathway.

Since the RPC and reductive citrate cycle (RCC) have the form of autocatalytic reaction sequences (Wächtershäuser, 1992), and their intermediates bear an invariant phosphate and carboxylate, respectively, both evidently originated as replication cycles, at different stages of evolution preceding the Last Common Ancestor (LCA). RPC synthesis of [PGA] and glyceraldehyde-3-phosphate [GAP], moreover, can be tied to initial steps in the emergence of life on Earth (Davis, 1999, 2006, 2012), as GAP is the phosphorylated derivative of an intermediate in

the spontaneous, formaldehyde-driven autocatalytic synthesis of glycoaldehyde, GO (Butelrow, 1861; Breslow, 1959). A transition from the 2- to 3-C self-propagating sugar, in this manner, implies continuity, at an elementary level, in the mechanism underlying the origin and evolution of life. And, while self-organization based on autocatalysis has long been central in theories on the origin of life (Hinshelwood, 1946; Calvin, 1969; Decker and Speidel, 1972), the standard model of evolution, drawn from the insights of Darwin (1859) and Wallace (1859), has resulted in pre-RNA developments, to the contrary, being attributed solely to chance (Research News: Max Planck Institute for Biophysical Chemistry, 2011).

A quantitative evaluation of the physicochemical forces driving competitive replication *in silico*, initially in the Qβ replicase system of Kramer et al. (1974), portrayed evolution as a damping response to distinct scalar forces within a non-equilibrium system (Davis, 1978a,b, 1996, 1998, 2000). With the realization that an interplay of physicochemical forces guided the trajectory of molecular and, by extension, biological evolution, outcomes beyond survival through natural selection, acting as a solitary agent of change, became foreseeable. Among them were stable and unstable states of coexistence between competing species, with non-linear rates of propagation, and episodic transitions, at variance with theories of gene propagation (Fisher, 1930) and polymer replication (Eigen, 1971) restricted to the maximization of rates. The generality of physicochemical forces also raised the possibility of bridging the conceptual gap between biological evolution, based on random mutations to a linearly encoded algorithm for self-propagation, and its purely chemical, prebiotic stages. Molecular evidence of pre-RNA, polycarbotide replicators (Davis, 2012, 2013b) offered a framework for reconstructing the transition.

Evidence presented here provides additional support for a pre-RNA, sugar-phosphate era in the origin of life. It derives mainly from the distribution of invariants among pyrimidine synthesis intermediates and transients in mainly ribozyme catalyzed reactions. Molecular models constructed to show progressively more advanced stages of polycarbotide replication are presented. They illustrate, in addition, the potential capacity of polycarbotides to fold into functional configurations required for stable tertiary structures and catalysis. The results presented broadly span an interval

from spontaneous autocatalytic synthesis of simple sugars from pre-sugar molecules, as broadly envisioned in the sugar model for the origin of life (Weber, 2001, 2007), to the RNA era. Replication is revealed to have evolved through a process of accretion, where a replicator at one stage served as a blueprint for development of its successor.

2. Ribonucleobase precursors

Nucleobases broadly form from three kinds of precursors (Fig. 1): (i) inorganics - CO_2, H_2O, HCO_3^-, and NH_3; (ii) amino acids – aspartate (Asp), glutamine (Gln), and glycine (Gly); and (iii) 1-C transfers from tetrahydrofolate (THF). Their synthesis also involves a ribose-5-P (RP) platform.

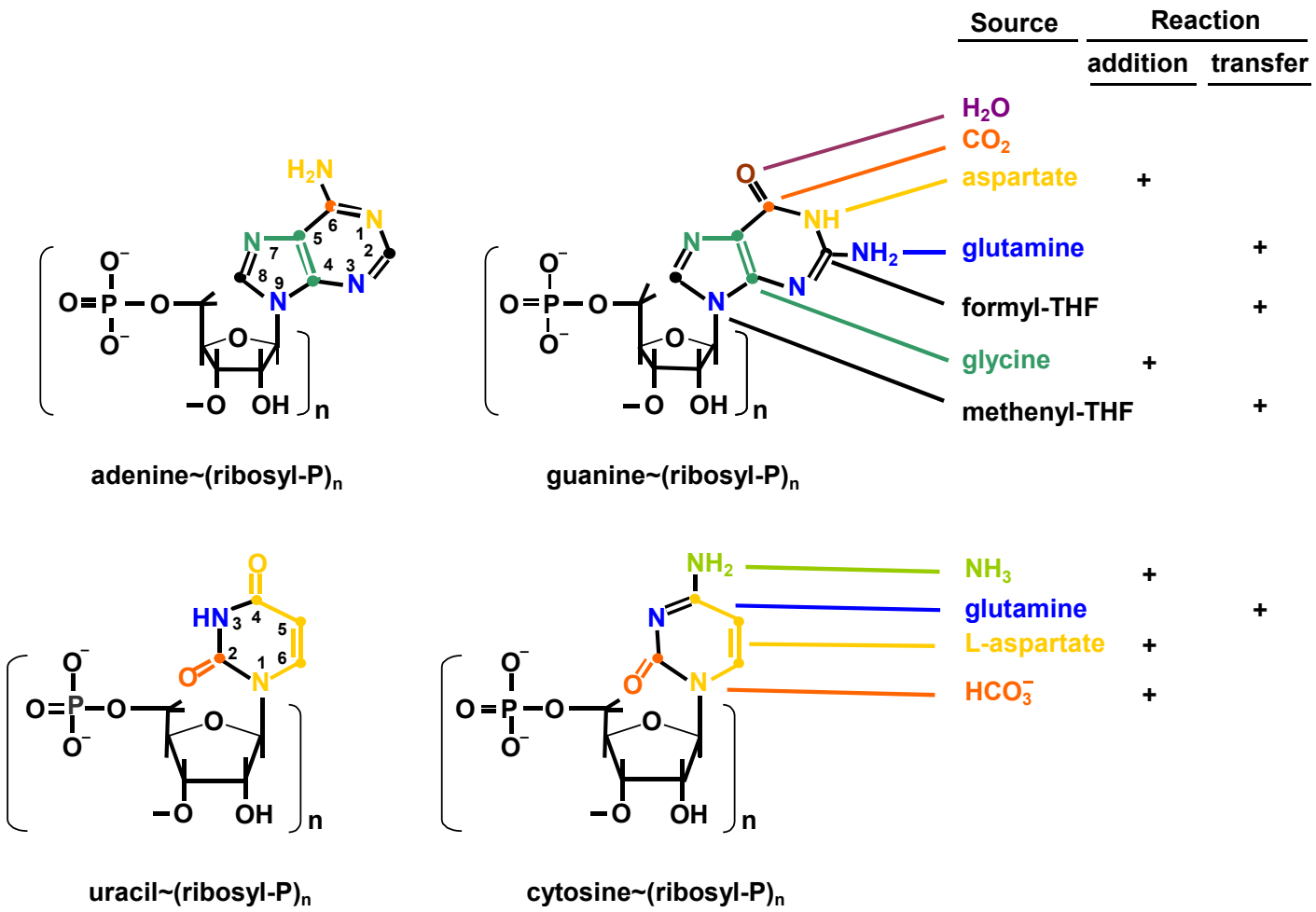


Fig. 1. Precursors to nucleobase synthesis. Amino acids Gln and Asp are N atom donors to purines and pyrimidines. A direct transfer of the Gln amide takes place in the synthesis of both kinds of nucleobases. Aspartate participates in an addition reaction. Non-N Asp atoms are deleted in the purine pathway. In the pyrimidine pathway, they are retained, apart from an α –carboxyl deleted in the step leading to uridine synthesis. Glycine contributes N and C atoms at purine sites 4, 5, 7. Formyl- and methenyl-group transfer from tetrahydrofolate (THF) respectively account for purine C2 and C8 (Wünschiers, 2012). Inorganic sources, CO_2, HCO_3^-, ammonia, and water contribute C, O, N, and H atoms, as shown. An invariant ribose-5-phosphate attached to purine and pyrimidine (beyond orotate) intermediates is depicted as an initial poly(pentosyl-phosphate) scaffold/cofactor.

High CO_2 levels in the early Archaean atmosphere and oceans (Walker, 1987), together with elevated ammonium in ancient clays of supra-crustal rocks (Honma, 1996), support nucleobase synthesis occurring by this stage of Earth history. Abiotic organics, including formate, form in the oceanic lithosphere, as it spreads outward from a tectonic ridge to off-axis hydrothermal seeps (Kelley et al. 2001; Proskurowiski et al., 2008; Lang et al., 2010) and subsequent subduction, make it a plausible location for the origin of life on Earth (Ward and Brownlee, 2000; Holm and Neubeck 2009) On the other hand, the synthesis of folates, such as THF, is GTP dependent (Schomburg, 2012). Thus, a simpler, pre-RNA 1-C donor, such as Gly (Piper et al., 2000), was the likely source of purine C-4 and C–5 in the nascent nucleobase synthesis pathway.

Aspartate, Gln, and Gly participation in nucleobase synthesis (Fig. 1) also places them in the pre-RNA era. This presupposes these amino acids had biochemical significance long before their incorporation into proteins. In this event, they were available for early entry into the genetic code. Biochemical evidence for their participation in pre-RNA metabolism thus could be conserved within the code. Figure 2 shows a mid-stage code, with 14 kinds of amino acids, reconstructed using the path-distance model (Davis, 2006). Codons for the nucleobase precursor amino acids can be seen to be located in regions encoding amino acids formed on short pathways. Here, the number of

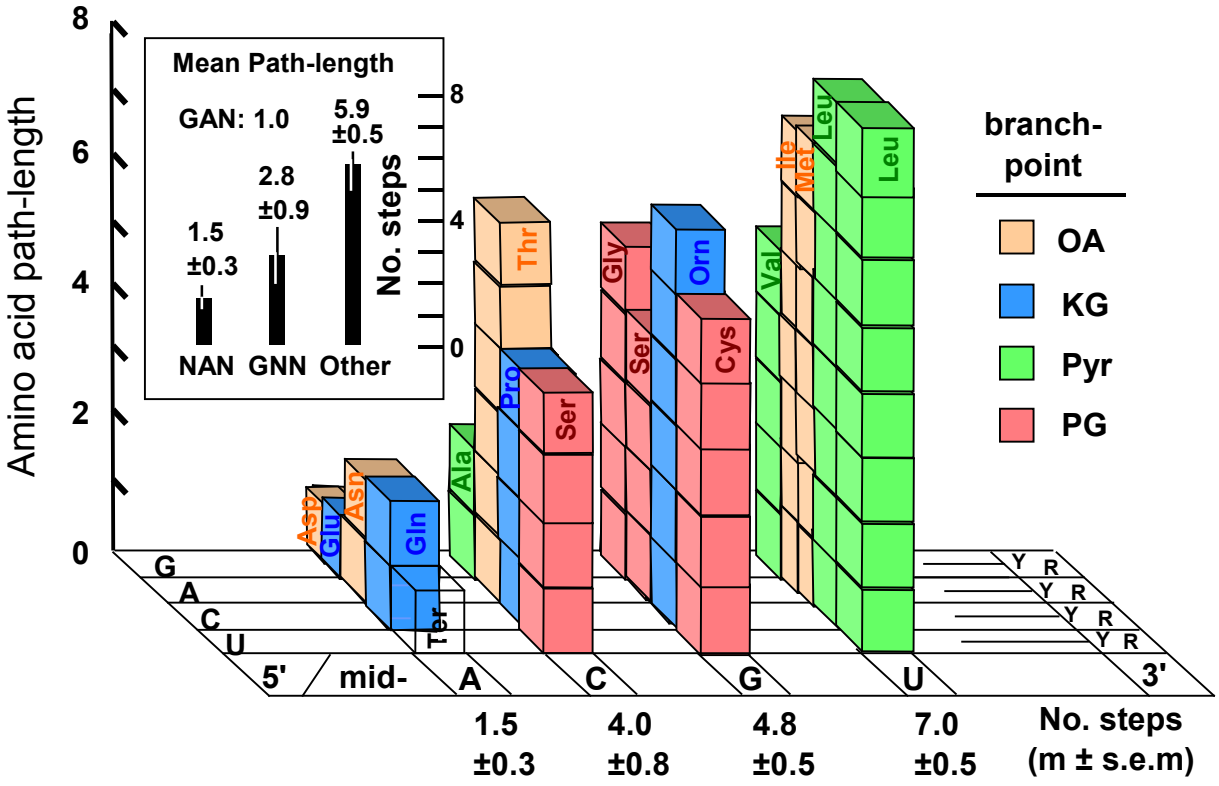


Fig. 2. Reconstructed mid-stage genetic code, with 14 amino acids (Davis, 2006). It shows triplets for nucleobase precursor amino acids, Asp[1], Gln[2], Gly[5], are in ancient codon clusters, characterized by 'short-path' amino acids. Aspartate[1] and Gln[2] are among four amino acids with NAN column triplets (inset) and they form on paths of 1.5 ± 0.3 steps (mean ± s.e.m.). Glycine[5] has codons in the GNN row, together with four other amino acids that form on paths of 2.8 ± 0.9 steps. Amino acids assigned triplets in the $\bar{G}\bar{A}$N ($\bar{A}$, not A; $\bar{G}$, not G) cluster have a significantly longer mean synthesis path of 5.9 ± 0.9 steps per codon set.

reaction steps required for synthesis is counted from the path branch-point in central metabolism, it serves as an indicator of amino acid time of entry into the code (Davis, 1999, 2002, 2006, 2009). Nucleobase precursors Asp[1] and Gln[2] (superscripts specify path-length), together with Asn[2] and Glu[1], have codons in the NAN set. They form in only 1.5 ± 0.3 (m ± s.e.m.) steps, on average. GNN row triplets encode Gly[5] and four other amino acids[1] that, collectively, form in 2.8 ± 0.9 steps.

[1] Glycine forms from serine in 5 steps; it can alternatively form in 2 or 3 steps by transamination of glyoxylate (product of RCC isocitrate) with an NH_2-donor amino acid (Nakada and Weinhouse, 1953).

GNN row triplets encode Gly[5] and four other amino acids[2] that, collectively, form in 2.8 ± 0.9 steps. The remaining 36 ḠĀN (Ḡ, not G row; Ā, not A column) in this mid-stage code were assigned as ten codon sets to seven amino acids[3]. These amino acids have significantly longer paths of 5.8 ± 0.5 steps ($p \leq 1.51 \times 10^{-2}$; one-way analysis of variance), indicative of their later entry into the code. Their triplets are also two mutations distant from GAN, while NAN and GNN triplets differ from GAN by no more than a mid- or 5'-base substitution. Significantly, GAN encodes Asp[1] and Glu[1] and they are the direct product of OA and KG amination, respectively.

Base triplet clusters, NAN and GNN, identified as primordial, overlap at GAN: a 4-set assigned to diacid homologues, Asp[1] and Glu[1] (Fig. 2). Both form in 1-step, on OA and KG amination, respectively. These amino acids, their amides, and carbamate were the apparent point of entry for N atoms, as ammonium, into ancient metabolic pathways, including those leading to nucleobase (Fig. 1) and α-amino acid synthesis (Davis, 2013a). Pathways branching from the citrate cycle, operating in the reverse direction under early Earth conditions (Wächtershäuser, 1992), produce nucleobase precursor amino acids. Invariant and nearly invariant terminal-carboxyls of RCC components conform with initial attachment to a pre-RNA scaffold/cofactor (Davis, 2013a). It is the vestige of a cofactor attachment site in metabolites within the RCC and its branching pathways that links this source of amino acids to the evolution of RNA and genetic code.

Base sequence identity among 'pre-species-divergence' tRNA established that amino acids from the same synthesis family charged phylogenetically related tRNA species (Davis, 2008, 2011). It is evident from this that tRNA had initially acted also as a cofactor in amino acid synthesis, in addition to its familiar role as an adaptor during translation. Early, dual function adaptor/cofactor tRNA molecules provided a mechanism for coordinating genetic code formation with the growth of amino acid synthesis pathways (Davis, 1999, 2006, 2008, 2009). Pre-divergence tRNA cognate with 'short-path' amino acids, alanine, Ala[2], and serine, Ser[4], displayed significant identity with $tRNA^{Asn}$ (Davis, 2013b). Since RCC metabolites and amino acid pathway intermediates alike bear an invariant carboxyl, pyruvate (Pyr) and PGA family amino acids - Ala[2], Val[5] and Ser[4], Gly[5], Cys[5],

[2] Includes ornithine, Orn[6], an α-amino acid intermediate of end-product/incorporated residue, arginine[9].

respectively - identified as precursors in pre-RNA metabolism, thus appeared to derive originally from OA, as an upstream precursor from Pyr and PGA, with a polyribotide cofactor. A 2'- or 3'-hydroxyl of the 3'-terminal ribose in the putative cofactor seemingly bonded with the α-carboxyl of an amino acid or intermediate, as in tRNA. The time of entry of a pre-RNA amino acid into the early code occurred no sooner than the tRNA takeover of an early polyribotide-dependent amino acid pathway. This, in turn, occurred no sooner than the inferred diversification of the cognate tRNA species from $tRNA^{Asn}$. Genetic code structure establishes that the time order of amino acid entry to the code correlates strongly with the number of post-'branch point' reaction steps required for synthesis (Davis, 1999, 2006, 2009, 2011).

Columnwise increments in mean post-'branch point' path-length, from 1.5 → 4.0 → 5.0 → 7.0 steps in the transition from NAN → NCN → NGN → NUN triplets (Fig. 2), show expansion from the initial N-fixers/donors code, encoding four amino acids (Asp^1, Glu^1, Asn^2, Gln^2), a STOP signal, with 16 NAN triplets, proceeded in a rather compact, error-minimizing manner (reducing lethal mutations to any of 48 nonsense triplets). With successive tRNA anticodon mid-base substitutions driving code expansion, amino acids sharing the same precursor (branch-point metabolite), and related adaptor/cofactor tRNA, acquired codons in the same code row; clarifying the source (Davis. 2006) of codon 5'-base invariance among synthetically related amino acids (Taylor and Coates, 1989).

$Aspartate^1$ and Gln^2 differ in the role they play in nucleobase synthesis and this exemplifies a range of disparities between them that has its origin long before the RNA era. In addition to different N-donor modes (Fig. 1), a wide difference exists in the number of kinds of amino acid species contributed to the code by Asp^1 versus Glu^1: 14 vs. 2 (Davis, 2013a). This outcome fits with Asp serving as a precursor, while Glu^1, and its amide Gln^2, were mainly N donors. A 200-fold faster rate of $Asp\text{-}tRNA^{Asn}$ and $Glu\text{-}tRNA^{Gln}$ amidation by Glu^1 (Sheppard, 2007), illustrated this. $Glutamine^2$ efficacy as an N donor, in fact, predates the appearance of enzymes (Nakada and Weinhouse, 1953). Distinct classes of synthetase charge $tRNA^{Asp}$ and $tRNA^{Glu}$ (Eriani et al.,), and Asp-tRNA-IA^{Asn} and Glu-tRNA-ID^{Gln} core structure groups[3] differ (Saks and Sampson, 1995; Davis, 2008).

[3] Roman numerals and upper case letters identify the tRNA core structure group.

Their frequency in fossil proteins also changes in opposite directions with phylogenetic depth (Brooks et al., 2002). From the spectrum of differences between Asp[1] and Glu[1], and their amides, the extra C atom in the Glu[1] side chain has had far-reaching consequences for evolution.

Table 1 lists pre-RNA amino acid precursors in nucleobase and coenzyme A (CoA) synthesis, and other related amino acids. They form on paths extending only 1 to 5 steps and, consistent with

Table 1. Triplet clusters encoding pre-RNA amino acids[1]

Amino Acids[2]	Pre-cursor	Triplet cluster[3]			Abiotic sources[4]	
		NAN	GNN	$\bar{G}\bar{A}N$	meteorite	discharge
		Nucleobase precursors + related amino acids				
Asp[1]	+	+	-	-	+(r)	+(r)
Asn[2]	-	+	-	-	-	-
Glu[1]	-	+	-	-	+(r)	+(r)
Gln[2]	+	+	-	-	-	-
Gly[5]	+	-	+	-	+(a)	+(a)
		Coenzyme A precursors + related amino acids				
Asp[1]	-	+	-	-	+(r)	+(r)
β -Ala[2]	-	-	-	-		
Val[5]	+	-	+	-	+(r)	+(r)
Ser[4]	-	-	-	+	-	+(r)
Cys[5]	+	-	-	+	-	NR

[1] Drawn from mid-stage genetic code (Davis, 1999, 2006).
[2] Superscripts denote number of post-'branch point' reaction steps in reconstructed tRNA-dependent synthesis pathway (Davis, 2013a).
[3] Overbars indicate excluded row (GNN) and column (NAN) in Fig. 2 code map.
[4] Based on results in Kvenvolden et al. (1970) and Miller and Orgel (1974) r, racemic mixture; a, achiral; NR, no sulfur-containing reagent.

.

expectation, acquired codons in the ancient NAN and GNN triplet clusters of the mid-stage code (Fig. 2). Non-encoded CoA precursor, β-Ala[2], results on decarboxylation of Asp[1], in a variation of an abiotic reaction (Doctor and Oro, 1972). Amino acids with chiral centers, from abiotic sources, are racemic, or nearly racemic (Breslow, 2011), in contrast to the homochirality of protein residues. This

strongly points to early involvement of self-organizing reaction mechanisms in the emergence of life, and it distances abiotic sources from this process.

3. Tandem invariants as evidence of pre-RNA pyrimidine cofactors

An Asp moiety α-carboxyl, acquired by pyrimidine synthesis intermediates in a reaction with carbomyl-P (Fig. 3), is retained unaltered until ortidine synthesis at step-4. Early intermediates in pyrimidine synthesis thus resemble those in amino acid synthesis pathways, where nearly all intermediates bear an invariant α-carboxyl. Genetic code structure and phylogenetics of tRNA species preceding the Last Common Ancestor (LCA) indicate that the α-carboxyl initially served as a tRNA cofactor attachment site, during synthesis of its cognate amino acid (Davis, 2008). The protein takeover of amino acid synthesis accounts for elimination of most tRNA cofactors. tRNA cofactors were virtually universal initially, from the ubiquity of amino acid synthesis intermediates bearing a free α-carboxyl group. It appears likely that the pre-RNA cofactor attached to Asp^{1} was a polyribotide forerunner of the tRNA scaffold, given the amino acid α-carboxyl bonds to the 2'- or 3'-hydroxyl of its 3'-terminal ribose.

A second pathway invariant, ribose-5-P (RP), can be seen to attach to uridine by a C1':N1 bond, at step-4. This signals attachment of a second cofactor. Decarboxylation eliminates the first invariant, together with its putative cofactor, in the next step of the reaction sequence. Attachment of RP provides direct evidence for a polyribotide cofactor in the nascent pyrimidine synthesis pathway. The RP platform in purine synthesis is likewise attributed to attachment of a pre-RNA, polycarbotide cofactor.

The cofactor exchange in pyrimidine synthesis resembles a 'hand-off' in a cofactor relay. A similar molecular mechanism, involving a tRNA cofactor exchange, has been found to account for anomalous triplet and tRNA assignments to Leu^{8} and Arg^{9} in the genetic code (Davis, 2013a). A shift in cofactor attachment site from uridine C6 to N1 occurs in the transition. Consequently, this shift accommodates formation of a double-helix, replicative-form of RNA and, significantly, of

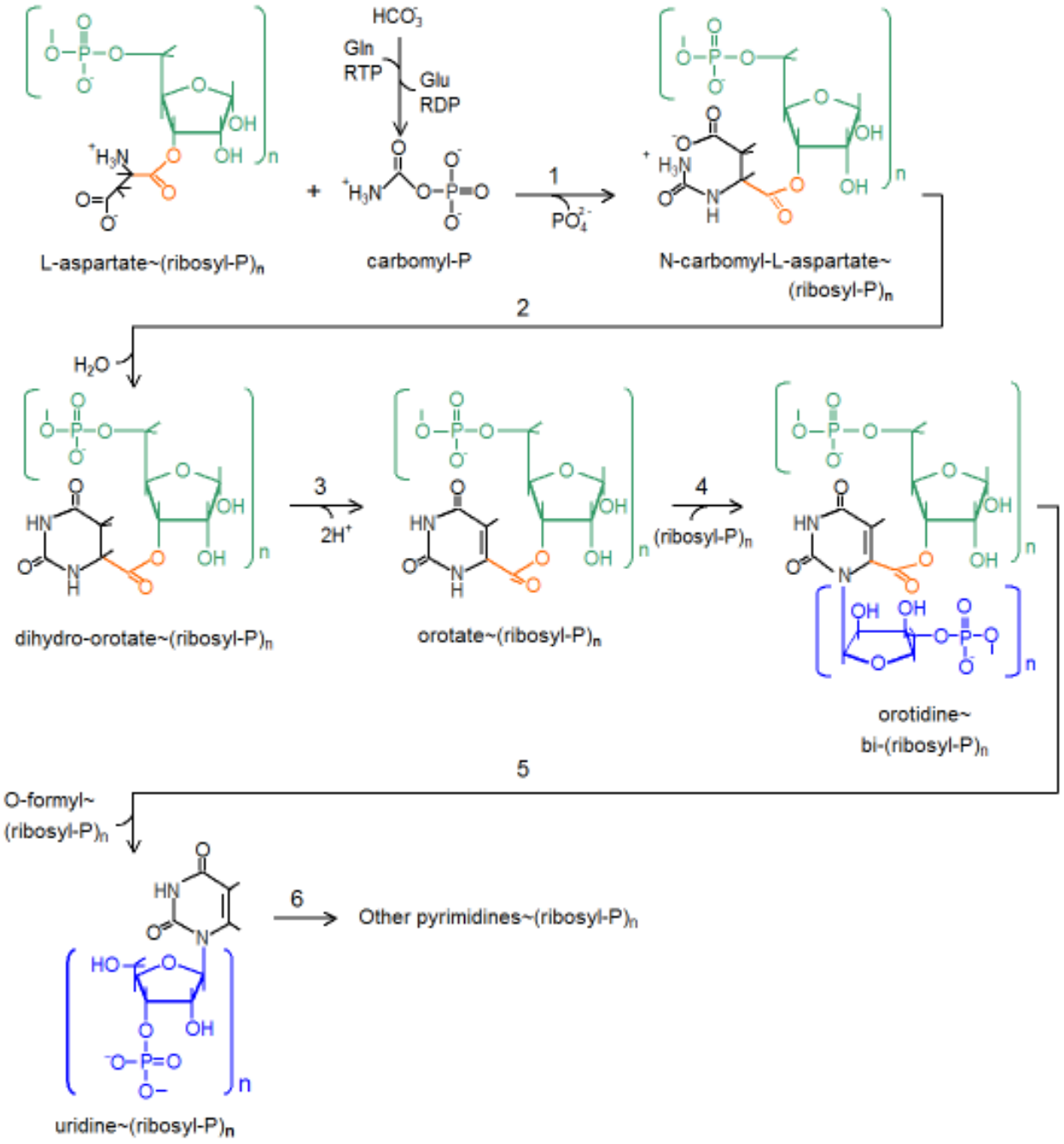


Fig. 3. Tandem invariants linked to a pre-RNA cofactor exchange in the pyrimidine synthesis pathway. Aspartate α-carboxyl is invariant from steps-1 to –4. At step-4, RP forms a C1:N1 bond with orotate and remains invariant thereafter. The α-carboxyl is deleted at the next step, resulting in a tandem distribution of path invariants. An exchange of pre-RNA cofactors, inferred to be the RNA polyribotide scaffold, accompanies the shift in cofactor attachment site to uridine N1. The exchange is thus conducive to double-helix formation. RDP and RTP designate ribosyl-di- and tri-phosphate. Orange, first invariant; green, proposed initial cofactor; blue, second invariant and cofactor

pre-RNA intermediates (Bernhardt and Sandwick, 2014).

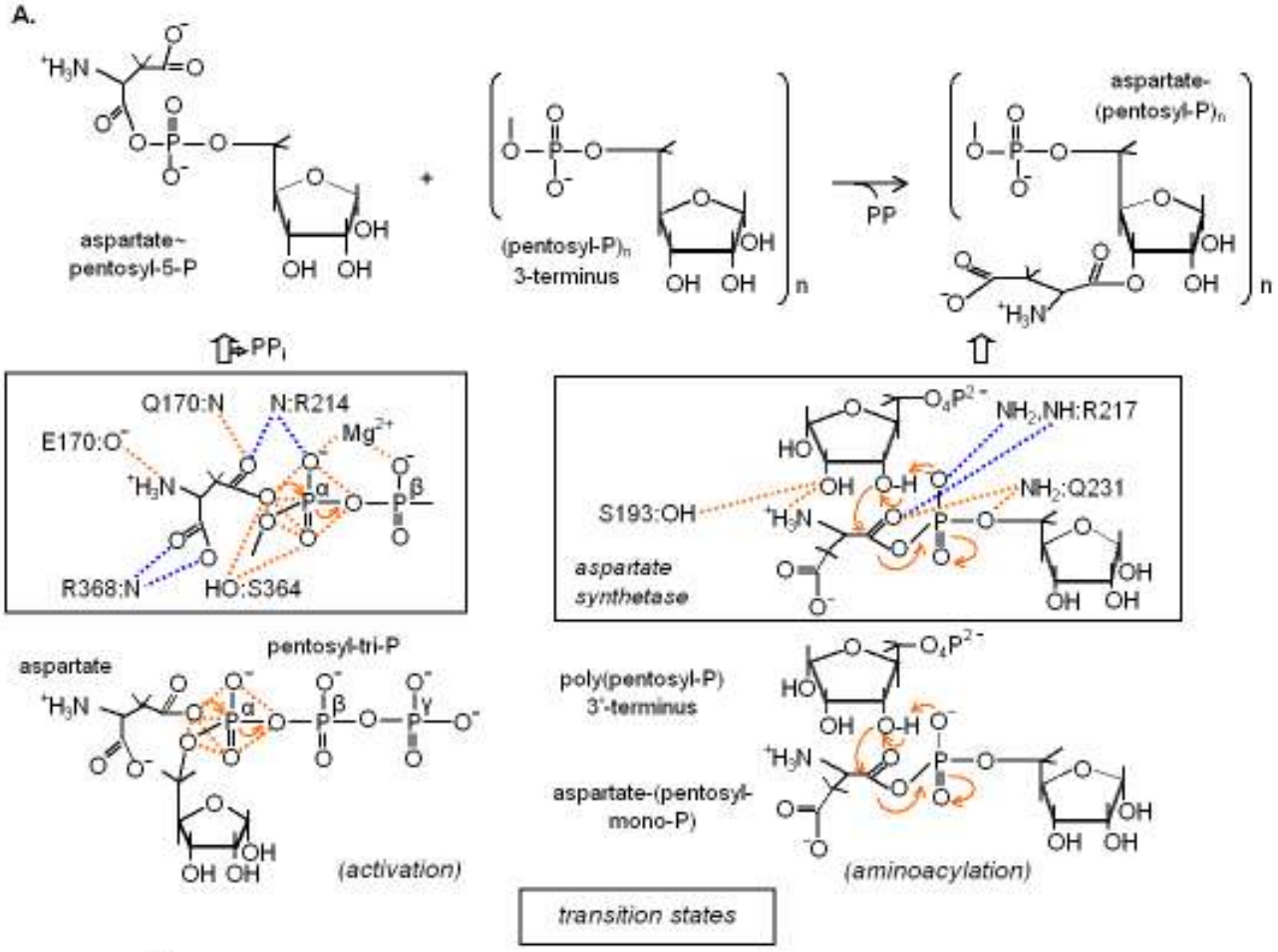
The cofactor exchange in pyrimidine synthesis resembles a hand-off in a cofactor relay. A similar molecular mechanism, involving a tRNA cofactor exchange, has been found to account for anomalous triplet and tRNA assignments to Leu^{8} and Arg^{9} in the genetic code (Davis, 2013a).

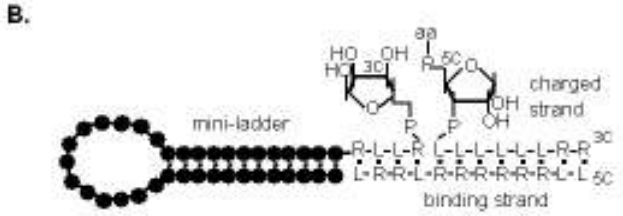
4. Carbotide era peptide synthesis

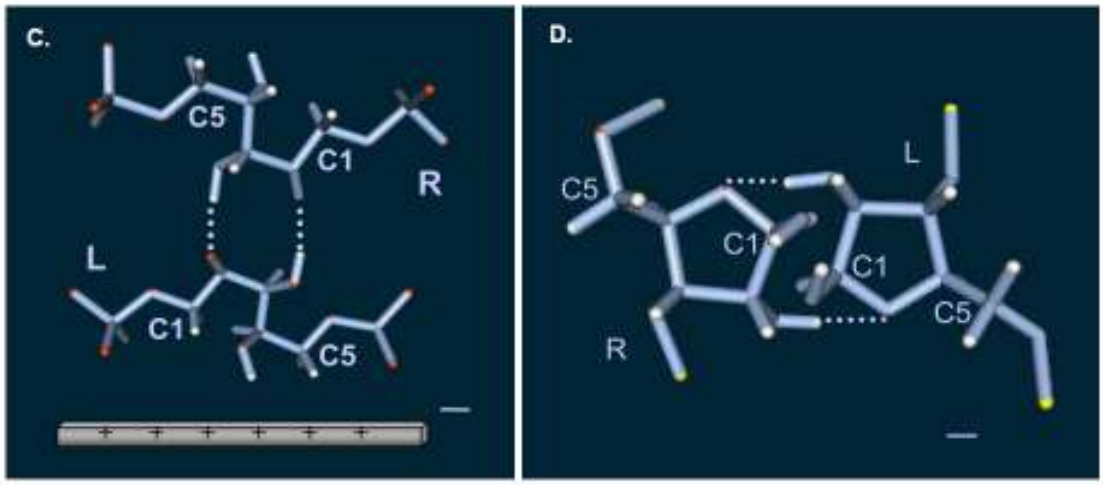
(i) *Amino acid activation and aminoacylation*

The participation of a small number of carbotide era amino acids in nucleobase synthesis (§3) raises the possibility that single-residue or random sequence oligopeptide were synthesized in the pre-RNA synthesis. A chain of 2 - 8 γ-linked glutamyl residues attached to THF (Schomburg, 2012), for instance, appears likely to have originated in the pre-RNA era, as Glu^{1}, together with Gln^{2}, predate the genetic code (§2). A-rich codons in the XAN (X, coding site) triplet cluster (Fig. 2) formed the first genetic code, in the path-distance model (Davis, 1999, 2006). This is consistent with formation of single residue or random (Asp,Glu,Asn,Gln) sequence oligopeptides in pre-code translation of a poly(A) template. Reaction paths in amino acid activation, cofactor acylation, and condensation provide additional evidence for pre-RNA peptide synthesis.

X-ray diffraction from the aspartyl-tRNA synthetase active-site of the archaeon thermophile *Pyrococcus kodakaraensis* at a resolution of 1.9 Å (Schmitt et al., 1998) indicate that amino acid activation proceeded via formation of a transient pentavalent α-P atom (Fig. 4A). It resolves into an activated amino acid (aminoacyl-ribosyl-5-monophosphate, in carbotide terms) with pyrophosphate (PP) release. There is no apparent nucleobase (adenine) participation in the reaction, supporting a pre-RNA origin of amino activation.







H-bonded L and R (surface-bound) D-ribulose-1,5-P pair

H-bonded L and R β-D-ribose-3,5-P pair

Fig. 4. Amino acid activation and cofactor aminoacylation with carbotide based reaction paths. **A.** High-resolution tRNA synthetase x-ray diffraction (Schmitt et al. 1998) indicates a transient penta-coordinated α-P atom of a ribotide-triphosphate bonds to the nucleophilic amino acid α-carboxyl O atom during activation. The transition state resolves into an aminoacyl-ribosyl-monophosphate and pyrophosphate molecule. No direct nucleobase participation is evident. Interactions stabilizing the transition state appear in the box above. Synthetase residues (*Pyrococcus kodakaraensis* numbering) and Mg^{2+} help neutralize the α-P group negative charge. Acylation of cofactor (poly(ribosyl-P)) with phosphorylated-Asp follows attack by a nucleophilic O atom, at ribose C3, on the activated amino acid carbonyl atom (α-carboxyl). Asp transfer to a carbotide cofactor results, as the transient penta-coordinated P atom bond with the amino acid is severed. Residue interactions stabilizing the acylation transition state derive from X-ray diffraction of *Escherichia coli* aspartyl-tRNA synthetase (Eiler et al. 1999). Orange lines, carbotide era interactions, include nucleobase precursors, a hydroxyl-bearing amino acid, and metal ion. Blue lines signify post-carbotide era interactions. Curved arrows show direction of electron flow. **B.** Tamura-Schimmel (2006) mini-helix RNA synthetase remodeled as a carbotide synthetase, comprising mini-ladder, binding, and charged strands, containing a binary sequence of H-bonded pentosyl-P monomers. **C.** Depicts putative H-bonds between counter-oriented (L, left; R, right) 2-keto-pentol (ribulose-1,5-bis-phosphate) monomers. C1 and C5 atoms are indicated, with the lower sugar-P bound, by charge attraction, to a cationic mineral surface. **D.** D-ribofuran monomer pair with O4..H-C2 bonds (2-4 bonding edge), with anti-parallel P groups at C3 and C5 (3-5 scaffold edge). These molecular models were made with Facio 18.1.1 (Suenaga, 2005) and optimized with the Gamess software package (Schmidt et al., 1993), based on restricted Hartree-Fock fragment molecular orbital calculations. White, H atom; red, O; yellow, P; C atoms not visible. Bar length, 1Å.

Transition state stabilization by synthetase residues and Mg^{2+} ions (Zurek et al., 2004) is critical for activation. Sugar-phosphate era interactions involving nucleobase precursor amino acids, or hydroxyl-bearing residue, contribute to stabilization of the transition state. Interactions with the positively charged Arg-214 and -368 guanidinium group occur (*P. kodakaraensis* residue numbers) A reconstructed carbotide era reaction path would require a metal ion, or substituted sugar, such as glucosamine, instead, to neutralize the negatively charged α-P group in the transition state.

High-resolution x-ray diffraction from the synthetase active site (Eiler et al., 1999) indicate aminoacylation follows an attack by a nucleophilic O3 atom of the 3-terminus ribose in the poly(RP)

scaffold of a tRNA cofactor, on an α-carbonyl atom in the activated amino acid (Fig. 4A). Transfer of an amino acid to its polycarbotide scaffold results, when the transient pentavalent P atom bond with Asp is severed. Synthetase interactions stabilizing the aminoacylation transition state are from the sugar-phosphate era, in the sense that they involve a nucleobase precursor or hydroxyl-bearing residue, apart from those with Arg-217 (*Escherichia coli* numbering). From the present perspective, this residue displaced an earlier, positively charged carbotide era molecule within the initial aminoacylation active site.

Tamura and Schimmel (2006) demonstrated aminoacylation of a mini-RNA helix under protein-free conditions, *in silico.* This followed incubation of a ternary complex comprising mini-helix, oligonucleotide bearing an activated amino acid, and binding-strand complementary to both mini-helix and charged oligonucleotide. With no evidence of direct nucleobase participation in the synthetase-catalyzed reaction (Fig. 4A), aminoacylation could be anticipated to proceed under carbotide era conditions. Figure 4B depicts a model carbotide synthetase. It substitutes a folded poly(RP) molecule, with a complementary H-bonded binary sequence, for the quaternary ribonucleobase sequence of the Tamura-Schimmel synthetase. Left (L) and right (R) oriented H-bonded ribulose-1,5-bisphosphate pairs, linked by a poly-P scaffold, illustrate an open configuration pentol monomer pair in Fig. 4C. A cyclic pentol-phosphate monomer pair is exemplified by H bonded β-D-ribose molecules in Fig. 4D. Anti-parallel P-groups, on the ‘3-5’ ribose edge, link monomers in each poly(β-D-ribosyl-phosphate) strand. H-bonds on the ‘O4 –HO2’ edge specifically bind L and R oriented ribose monomers (Fig. 4D). It is noteworthy that the double helix utilizes the ‘3-5’ pentose edge of its poly(β-D-pentosyl-phosphate) scaffold for phosphorylation, since this leaves the pentose ‘O4-HO2’ edge, in the proposed carbozyme, available for folding through H-bond formation.

(ii) *Amino acid condensation*

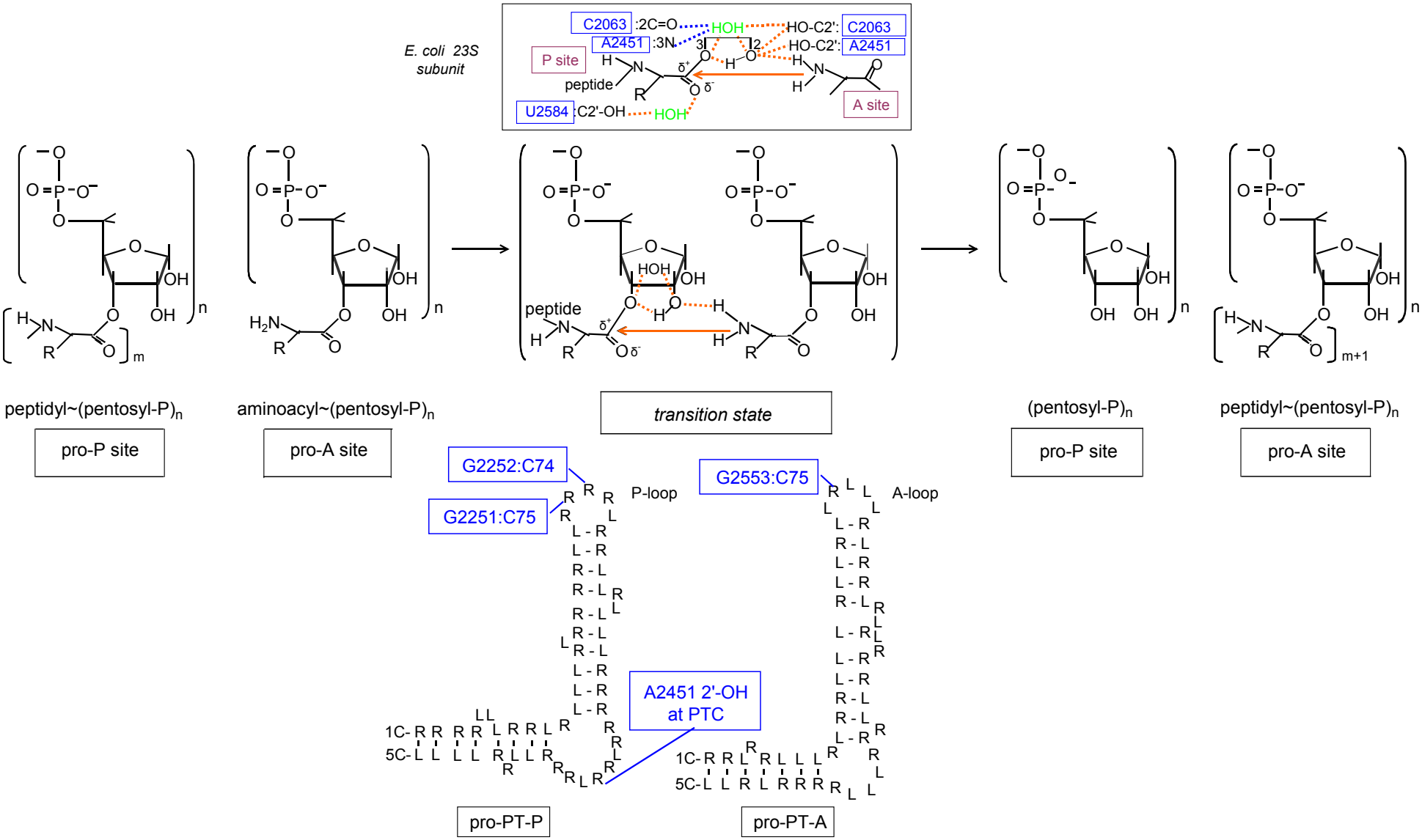
Polypeptide chain extension at the peptidyl transfer (PT) center, in the large ribosomal subunit, entails a transient proton shuttle between the incoming amino acid α-amine, positively polarized

carbonyl C atom, in the terminal polypeptide residue, and ribose O3 atom in its cofactor scaffold (Fig. 5). Formation of a transient H-bond between O atoms at ribose C2, in the P-site cofactor, and incoming amino acid carboxyl catalyzes amino acid condensation (Gindulyte et al., 2007). H -bonds between ribose C2 hydroxyl in nucleobases C2063, A2451, and U2584 (*E. coli* numbering) and terminal ribose O2 atom of the P-site cofactor, or screening water molecule, stabilize the transition state (Rodnina et al. 2007). In addition, C2063 and A2451 nucleobases H-bond with an active-site water molecule. On substituting them with equivalent polycarbotide interactions, the reaction path in peptide bond synthesis is seen to conform with a sugar-phosphate era process (Fig. 5).

Phylogenetic analysis of ribosome structure, at atomic resolution, revealed ribosome evolution proceeded by accretion, commencing at the universal PT center (Petrov et al., 2014). A symmetry in the (protein-free) active site led Agmon and her associates (Agmon et al., 2005, 2009; Davidovich et al., 2009) to propose that 'pre-proto-ribosome' A and P site ribozymes, with a shared stem-loop-stem configuration, cooperatively catalyzed initial peptide synthesis. Identification of this symmetry agreed with an A and P site symmetry (Davis, 1999) inferred from initiation codon 5'-base

---

Fig. 5. Pre-RNA peptide bond formation. A transient proton shuttle is shown to form the peptide bond. The reaction is initiated when a proton from O2 is donated to its vicinyl (leaving) O3 atom in the terminal ribose of a polycarbotide cofactor at the putative carbosome P-site. There is a simultaneous transfer of a proton from the α-amino group of an incoming amino acid at the A site. An attack by its nucleophilic NH_2 group on the positively polarized ester carbon forms the peptide bond, to extend the homologous or, conceivably, pre-RNA code heterologous polypeptide chain (Trobro and Aqvist, 2005; Rodnina et al, 2006; Gindulyte at al., 2007). Hydroxyl formation at C3, in the leaving cofactor ribose, accompanied by incorporation of the new amino acid with concurrent translocation of the poly(ribotide) cofactor from A to P sites. Orange and blue lines respectively mark pre- and post-carbotide era interactions, as in Fig. 4. Nucleobase sites (*E. coli* numbering) mark stabilizing interactions. Ancestral A- and P-site peptidyl transferase (PT) ribozymes identified by Agmon et al. (2009) and Davidovich et al. (2009) have been remodeled as carbozymes, with known PT:tRNA interaction sites shown.


E. coli 23S subunit
C2063 :2C=O
A2451 :3N
U2584 :C2'-OH
HOH
HO-C2': C2063
HO-C2': A2451
P site
A site
peptide
peptidyl~(pentosyl-P)n
pro-P site
aminoacyl~(pentosyl-P)n
pro-A site
transition state
(pentosyl-P)n
pro-P site
peptidyl~(pentosyl-P)n
pro-A site
G2252:C74
G2251:C75
G2553:C75
P-loop
A-loop
A2451 2'-OH at PTC
pro-PT-P
pro-PT-A

degeneracy analogous to internal codon 3'-base degeneracy (Crick, 1966). Folded poly(pentosyl-phosphate) carbozymes (Fig. 4C, 4D), similar to those reconstructed from A and P site ribozymes (Fig. 5), according to these findings, catalyzed condensation of nucleobase precursor amino acids (§3) in the pre-RNA era.

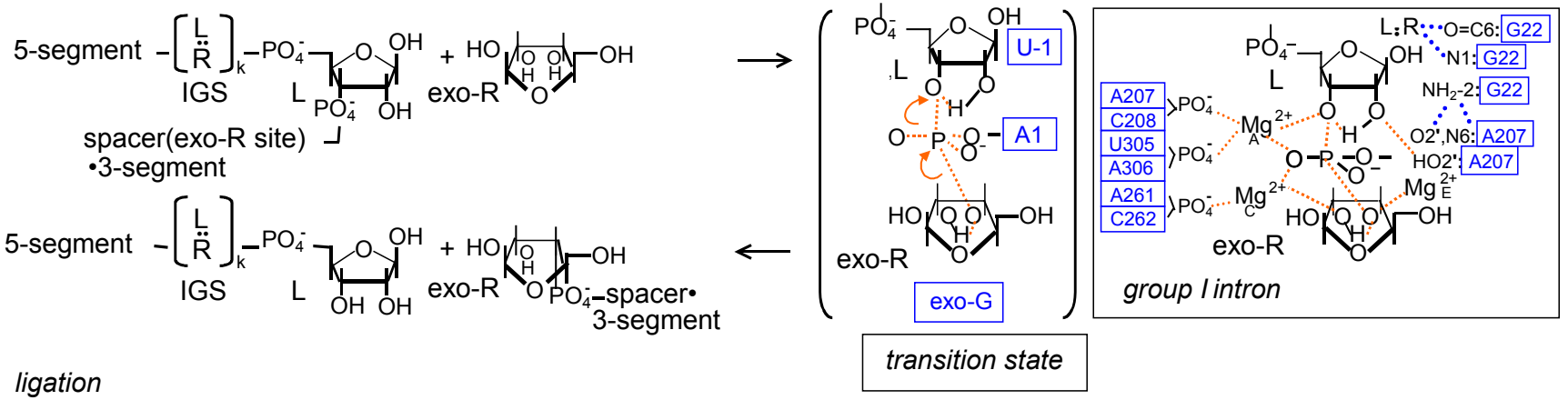
5. Ribozymes as a blueprint for carbozyme catalysis

Figure 6A(i) and (ii) depict cleavage and ligation steps in the self-excision of a spacer sequence within a poly(RP) strand, based on the group I intron mechanism. Both reactions proceed through formation of a transient pentavalent P atom, with coordinately bonded active site (Mg^{2+}_A, Mg^{2+}_C) and peripheral (Mg^{2+}_E) metal ions stabilizing the transition state (Stahley and Strobel, 2005; Cochrane et al., 2007; Forconi et al., 2007). With a carbotide vintage reaction path in this process, and equally early paths characterizing the activation and acylation steps of peptide synthesis (§4), the pre-RNA era had significant molecular diversity. This does not imply, however, that the group I intron preceded RNA, or the LCA (Haugen et al., 2005). Intron splicing mechanisms of various kinds are, however, widely distributed among different phyla, raising the possibility of an early origin of some form of intron splicing and, possibly, lateral transfer. As the group I intron utilizes a mechanism of cis-acting pre-protein catalysis, it illustrates a form of pre-cell compartmentalization, with fundamental implications for the origin of life.

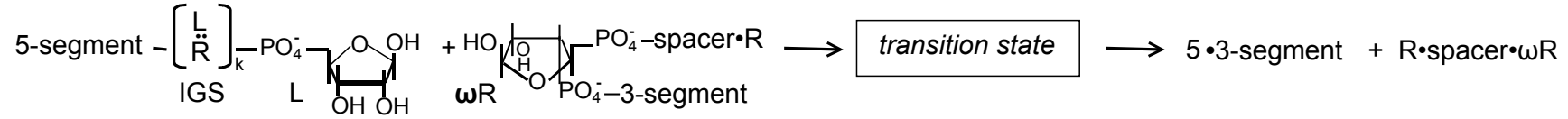
Riboswitch responsiveness to the product of an encoded enzyme, exemplified by *glmS*, a glucosamine-6-phosphate (GlcN6P) binding ribozyme, located in an untranslated region of the GlcN6P synthetase gene transcript, expands the range of ribozyme functions (Winkler et al., 2004). It also enlarges the potential scope of carbozymes. Figure 6B depicts the GlcN6P dependent autolysis of a poly(RP) strand by a carbotide model *glmS*. Molecular dynamical and force field calculations revealed an ammonium group in the glmS active site (Banas et al., 2010) served as a general acid during catalysis, donating a proton to O5 in the leaving ribose (O5'@G1 in *glmS*). Involvement of a proton donor in the *glmS* catalyzed reaction distinguishes it from the foregoing

A. Self-splicing

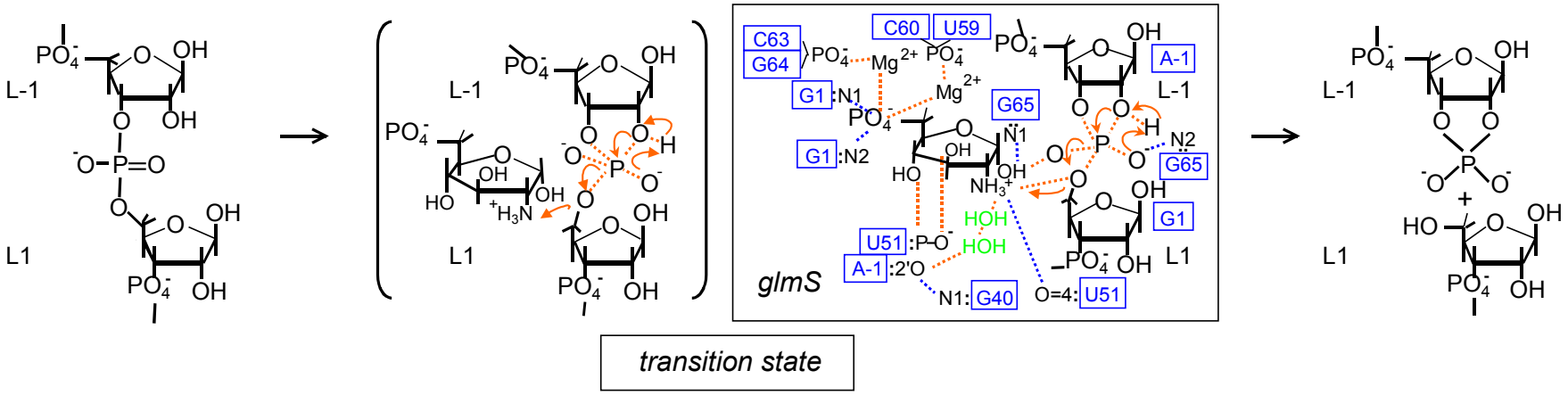
(i) *cleavage*



(ii) *ligation*



B. Cofactor-bearing



C. Self-replicating

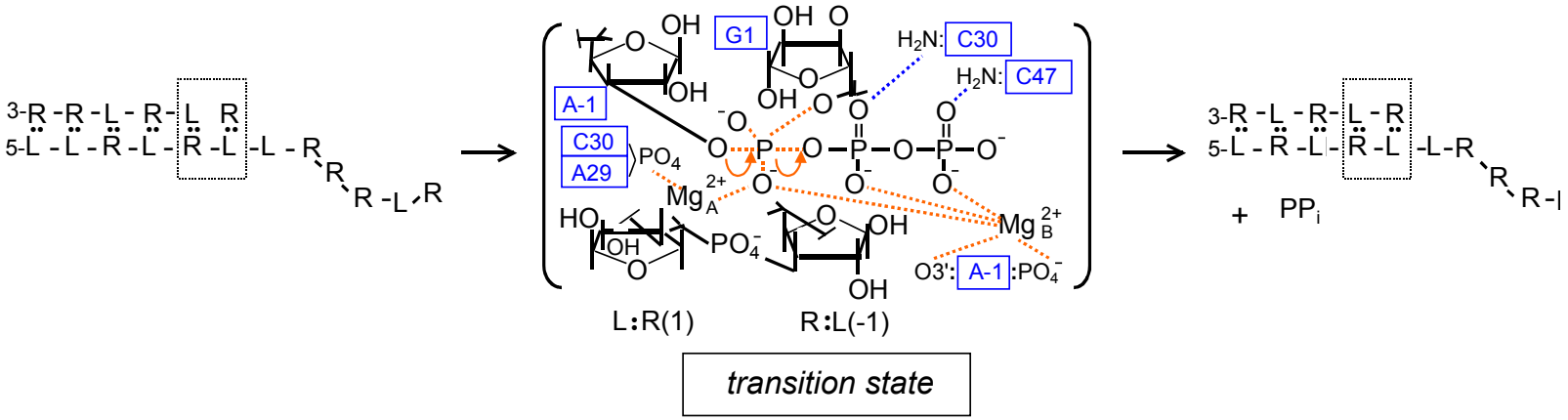


Fig. 6. Ribozyme and pre-ribozyme catalysis. **A.** Self-splicing of a binary carbotide strand based on the group I intron. Spacer-segment excision within a strand of complementary H-bonded left (L) and right (R) oriented ribosyl-P monomers (see Fig. 4d). (i) Cleavage at the 5-segment•spacer junction (5'-exon•intron junction in group I intron) on alignment on an internal guide sequence (IGS). A transient penta-coordinate P atom forms at the L-1/R1 link (U-1/A in RNA; indicated by blue letters), as in *Azoarcus sp.* pre-tRNA^{Ile} intron active site (Stahley and Strobel, 2005). Transesterification links 3'-hydroxyl of bound exogenous R-oriented ribose (exo-G) to 5-terminus of a spacer•3-segment (intron•3'-exon), freeing HO3 at 5-segment 3-terminus. (ii) 5- and 3-segment ligation follows attachment of latter 5-phosphate to former ribose O3, releasing spacer and terminal R ribose (intron ωG) via a transient penta-coordinated-P. *Tetrahymena* group I intron active site nucleotides (Hougland et al. 2006) are shown. **B.** Glucosamine-6-phosphate bearing carbozyme self-cleavage, as in *T. tercongenesis glmS* riboswitch (Klein et al., 2007). Curved arrows depict electron flows. Cofactor amide contributes a proton in C5 hydroxyl formation (Banas et al., 2010) in downstream ribose (at *glmS* G1) at cleavage site. A ribosyl cyclic mono-phosphate forms upstream (at A-1 in *glmS*). Orange and blue broken lines apply as in Fig. 4. Magnesium^{2+} ion interactions are based on *Bacillus anthracis* active site (Cochrane et al., 2007). **C.** Pre-RNA replication is based on that by artificial class I RNA ligase, *in silico* (Attwater et al., 2013). Ligase active site studies (Shechner et al., 2009; Sgrignani and Magistrato, 2010) furnish evidence for interactions shown. Non-bonding O atoms of G1 β- and γ-phosphoryl groups H bond to C30 and C47 amide, respectively, with direct effect on ribozyme catalytic efficiency.

carbotide cleavage reactions (Fig. 4A, 6A). Cofactor capacity to produce a novel, efficient reaction path is also apparent. The *glmS* reaction path, nonetheless, remains of carbotide vintage, for the ribose 2-OH, of A-1, initiates the reaction with a nucleophilic attack on a non-bonded O atom of the scissile phosphate, to produce a ribosyl cyclic phosphate. Ribozyme nucleobases G1, G65, and G40 (*Thermoanaerobacter tencongenesis* numbering) play a key role in positioning both cofactor and 2'OH group in an in-line configuration conducive to cleavage (Klein et al. 2007). Any pre-RNA model of the *glmS* ribozyme will require equivalent carbotide binding interactions.

Obtaining an L1-ligase derived ribozyme polymerase (202 nt.) capable of replicating an RNA strand longer than itself (Attwater et al., 2013) established that RNA could, in principle, self-replicate prior to the appearance of proteins. Figure 6c depicts a carbotide model of the L1-ligase active site with relevant nucleobases and metal ions identified by Sgrignani and Magistrato (2012), in a quantum mechanical/molecular dynamical analysis of the active site structure identified by Shechner et al. (2009). Extension of a poly(RP) strand by condensation with an RTP molecule is an intrinsically carbotide process. Amides of ligase nucleobases C30 and C47, however, stabilize the transition state through H bonds to non-bonding O atoms of the β and γ phosphates. A carbotide era polymerase would plainly need to provide alternatives to the nucleobase proton donors.

6. Discussion

6.1 Pre-RNA pathway invariants as polycarbotide cofactor attachment sites

The pattern of invariants among nucleobase intermediates (§3) parallels that found in amino acid synthesis pathways (Davis, 2008, 2013). Thus, the role of tRNA in contributing to α-carboxyl ubiquity within amino acid synthesis pathways provided a framework for interpreting similar features in nascent nucleobase pathways. In particular, an invariant ribotide in the pathways for purine and post-orotate pyrimidine synthesis duplicated α-carboxyl invariance among amino acid intermediates. By extension, a polyribotide cofactor could be inferred to have participated in pre-RNA nucleobase synthesis.

Prior to attachment of RP to ortidine, at step 4, pyrimidine synthesis intermediates were noted to

retain an Asp α-carboxyl (Fig. 3). Hence, orotidine had dual invariants: a ribotide at N1 and an earlier α-carboxyl at C6. Dual invariants are also a feature of Leu^{8} and Arg^{9} synthesis: α-isopropyl-malate (step-5, Leu^{8} path) and arginine-succinate (step-8, Arg^{9} path). Significantly, both amino acids have tRNA and codons unlike those of synthetically related amino acids: tRNA-II^{Leu} isospecies read UUR, CUN vs. tRNA-IA^{Ala} GCN, tRNA-IA'^{Val} GUN, while tRNA-IA^{Arg} isospecies read AGR, CGN vs. tRNA-ID^{Gln} CAR, tRNA-ID^{Pro} CCN, tRNA-ID^{His} CAY. Elimination of the initial invariant α-carboxyl in both nascent tRNA-dependent amino acid synthesis pathways, combined with retention of the second invariant α-carboxyl, directly conformed with an exchange of tRNA cofactors. This exchange, at a mid-point in the sequence of reactions leading to synthesis, provided an explanation for their anomalous cofactor and codon assignments (Davis, 2013a). Based on these findings, α-carboxyl deletion at step-5 in pyrimidine synthesis, combined with ribotide retention, conforms with the exchange of a pre-RNA cofactor. As the initial pre-RNA cofactor was attached at orotidine C6, discarding it, while retaining the second cofactor at N1, shifted the cofactor attachment site (C6 → N1). This shift resulted in a change compatible with poly(nucleobase/ribotide-scaffold) double-helix formation. The putative cofactor exchange is thus seen to be a step toward forming the (pre)polynucleotide replicative-form. Purine synthesis on an RP platform (vestige of polyribotide cofactor/scaffold), attached at N9, provides added support for this interpretation.

6.2 Design features of polycarbotides

Beyond substantiating the presence of pre-RNA cofactors in nascent nucleobase pathways, parallels between the distribution of invariants in nucleobase and amino acid synthesis left early cofactor structure and function largely unresolved. Occurrence of a ribotide platform in purine and pyrimidine synthesis makes it likely that the Asp moiety cofactor, in pyrimidine synthesis, was a forerunner of the tRNA polyribotide scaffold. This raises a question concerning the nature of identity elements within a polycarbotide cofactor. Recognition of Asp-$tRNA^{Asn}$ and Glu-$tRNA^{Gln}$ by a macromolecular ribonucleoprotein transamidosome (Blaise et al. 2010), responsible for amidation of Asp^{1} or Glu^{1} in each misacylated tRNA species, derives from specific nucleobases in the Asp-

$tRNA^{Asn}$ and Glu-$tRNA^{Gln}$ sequence (Bailly et al., 2006). Thus, the polyribotide cofactor attached to Asp (Fig. 3) evidently contained identity elements to recruit pyrimidine-specific carbozymes. This introduces the first carbotide design feature:

Complex monomer sequence

Structural and catalytic RNA molecules self-assemble into folded configurations stabilized by H bonds on the Watson-Crick, Hoogsteen, and sugar edge (Leontis and Westhof, 2001;Tamura and Holbrook, 2002), in addition to coordinate bonds with divalent metal ions. Naked RNA zipper interactions employ sugar-to-sugar H bonds, consistent with pre-RNA information bearing sequences in polycarbotides. As in RNA, furthermore, functional polycarbotides necessarily incorporate a second design feature:

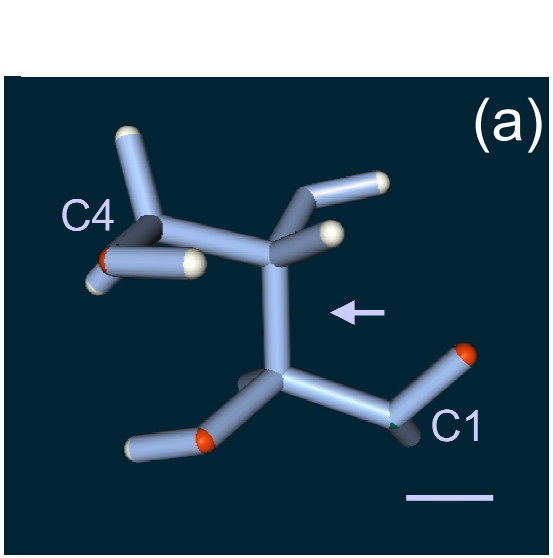
Functional tertiary structure

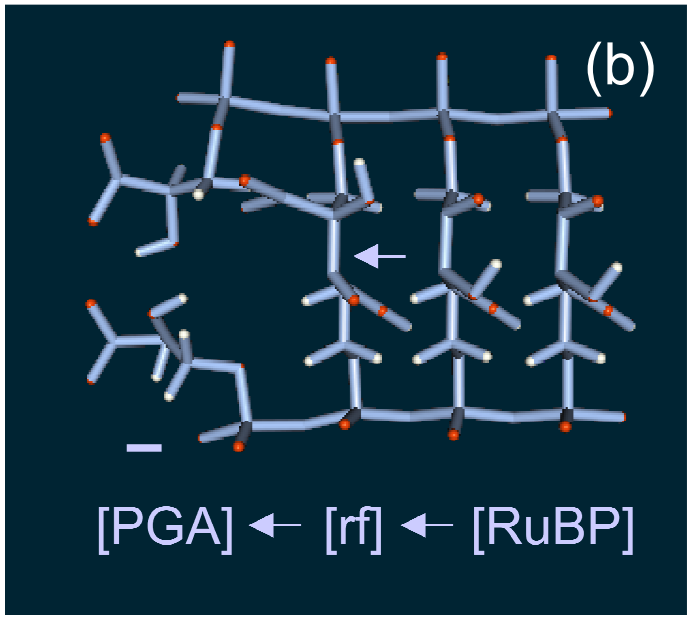
Carbotide-centered transition states in ribozyme-catalyzed reactions (Sponer at al., 2012) were seen to span activation in peptide synthesis to ligase condensation during RNA replication (§4, 6). Carbocentricity in these reactions supports polycarbotide catalysis, contingent upon self-annealing into a functional structure. It also implies that a broad range of reactions originated in the pre-RNA, sugar-P era. For evolution from carbozyme to ribozyme catalyzed reactions, carbozymes, like ribozymes, required a third feature:

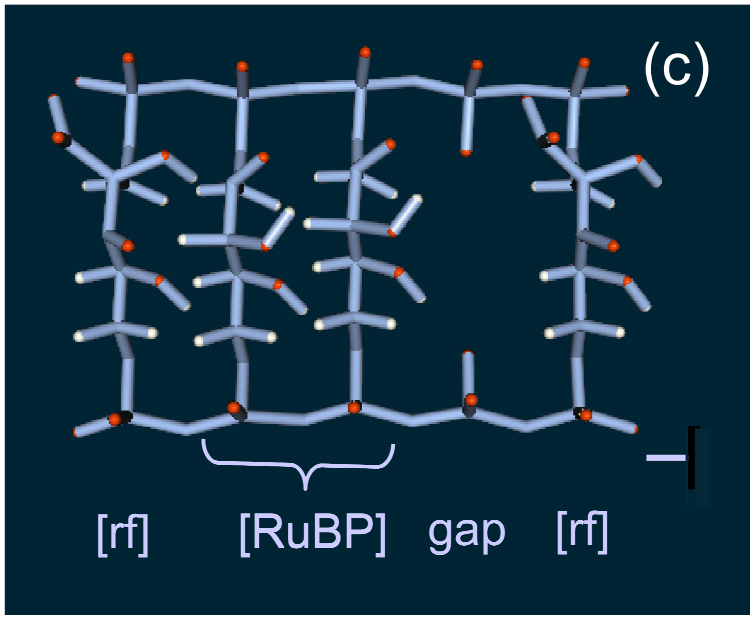
A replicative-form

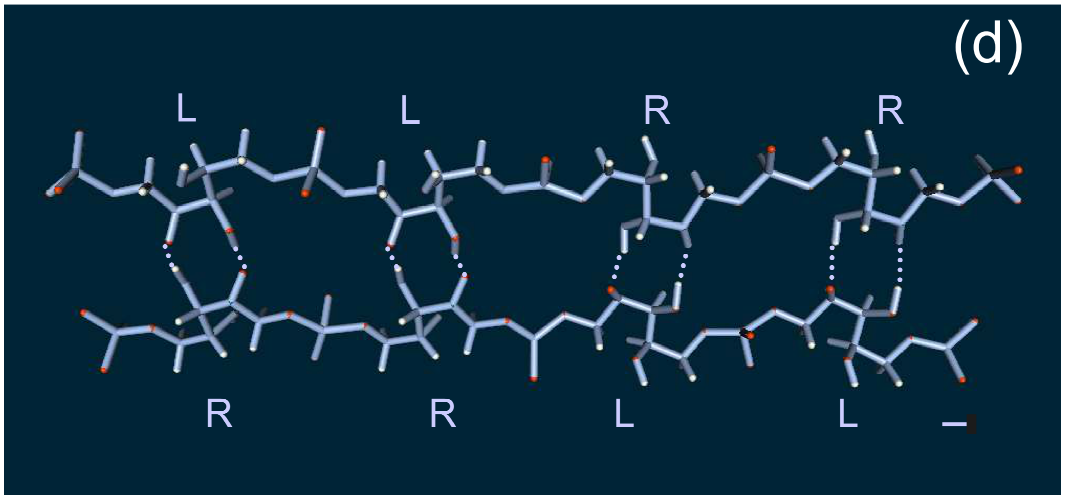
Notwithstanding their capacity to form functional tertiary structures, protein residue sequences plainly lack a replicative-form, consistent with the post-RNA origin of protein synthesis. Synthesis of a simple pre-RNA oligopeptide, for this reason, would the require encoding by an early self-propagating polycarbotide.

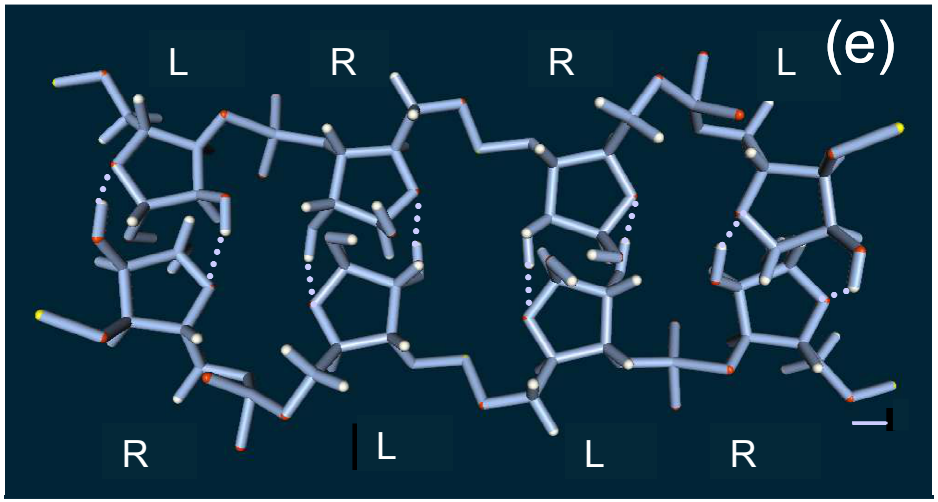
Figure 7 shows a set of increasingly advanced model polycarbotide self-propagators. Indicative of

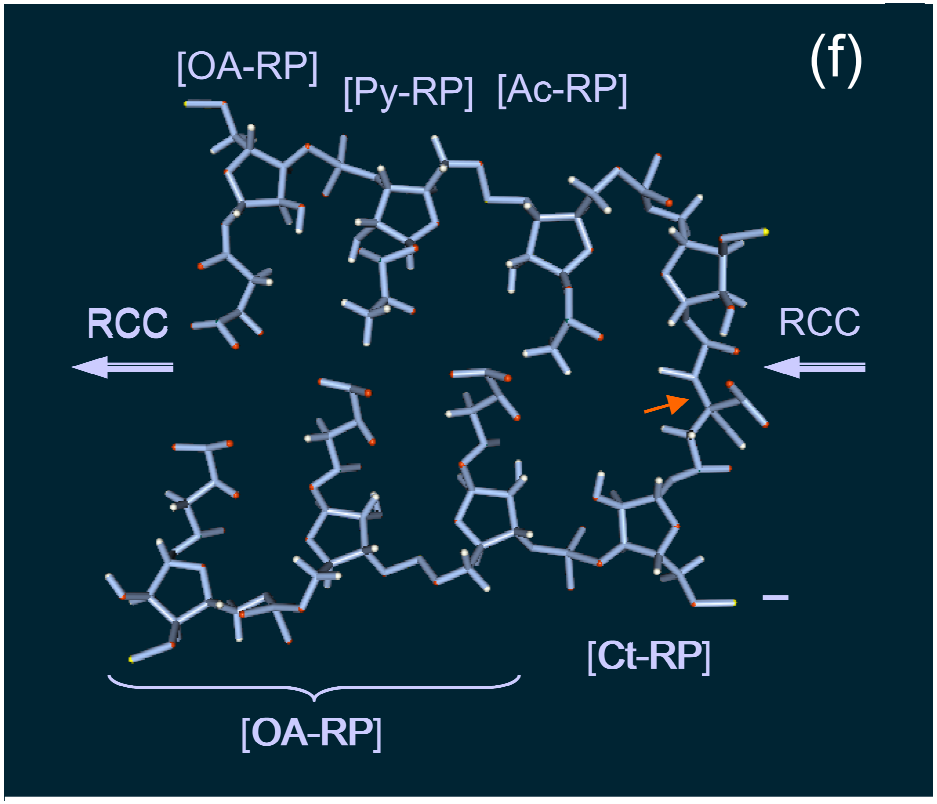

(a)
C4
C1


(b)
[PGA] ← [rf] ← [RuBP]


(c)
[rf] [RuBP] gap [rf]


(d)
L
L
R
R
R
R
L
L


(e)
L
R
R
L
R
L
L
R



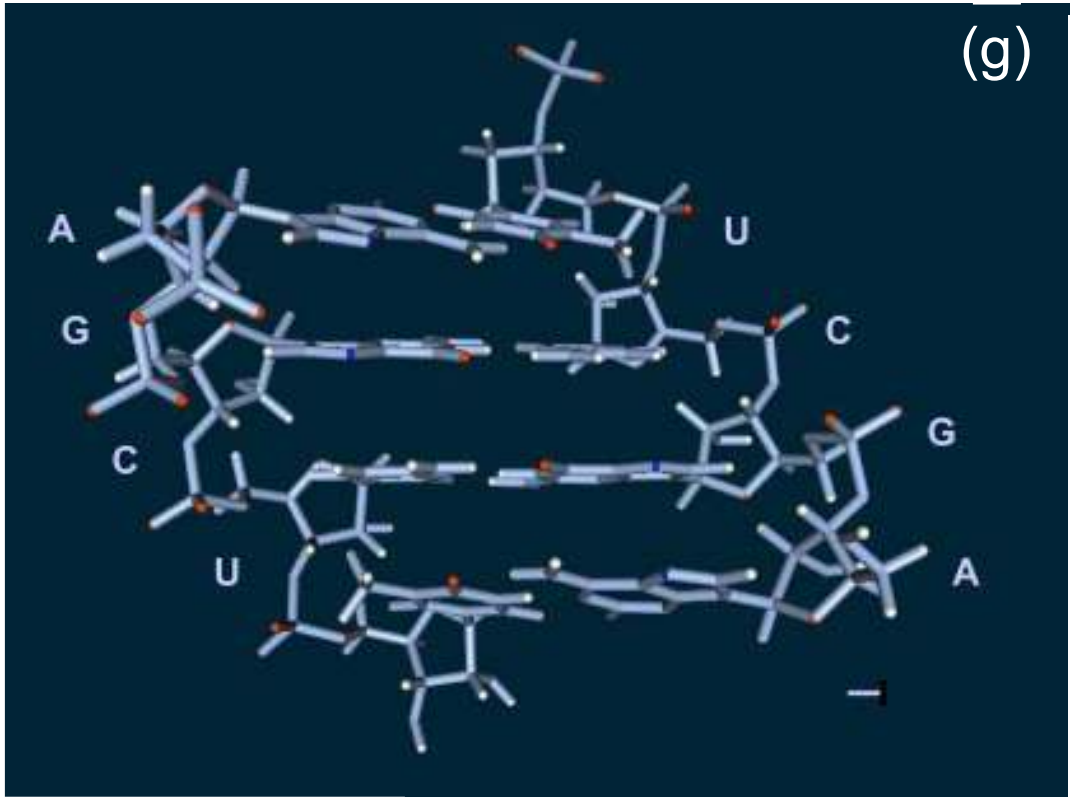


Fig. 7. Evolution of replication based on the distribution of path invariants. (a) Aldotetrose cleavage between C2 and C3 (arrow), in a spontaneous retro-aldol formose reaction, yields two molecules of a 2C self-propagating sugar, GO. (b) Ladder-like molecule formed with a metabolite in autocatalytic RPC, pentose RuBP, bearing invariant phosphates at each terminus. Carboxylation at C2 produces CKAP - the replicative form [rf] of the RPC polymer. A retro-aldol type reaction cleaves CKABP between C2 and C3 to yield two PGA monomers. (c) A binary RuBP/gap sequence attached to a poly(P) scaffold forms a transmissible encoding mechanism. (d) An RuBP ladder with stereo-specific H bonds (O2••H-O3) between complementary strands of anti-parallel keto-pentol pairs, oriented left (L) and right (R). Replication of this duplex severs comparatively low energy H-bonds. (e) β-D-Ribose, formed by ribulose cyclization, H bonded (O2-H••O4) on its 2-4 edge with a counter-oriented ribose monomer, in complementary anti-parallel polyribotide chains, linked by 3-5 edge phosphate bonds. (f) Autocatalytic RCC intermediates, including Citrate (Cit) and OA, have an invariant carboxyl at each terminus, indicative of parallel polyribotide scaffolds; Co A likewise has an invariant ribosyl-P in its synthesis pathway (§7.3). A polymeric form of the RCC shows Cit cleavage at C2-C3 (orange arrow) leads to two OA, making it a 4C self-propagator. (g) Type A RNA with H bonded, purine:pyrimidine base pairs on anti-parallel β-D-ribose-P scaffolds; the pentotide is an invariant in nucleobase synthesis. Each model replicator was constructed as specified in Fig. 4.

pre-RNA polycarbotide replication (Davis, 2012), each terminus of RPC intermediates contains an invariant phosphate. In addition, RuBP cleavage produces two PGA molecules, with a new PGA synthesized every three cycle rotations, making the RPC a CO_2 driven autocatalytic cycle and, at polymer scale, a replication cycle. As RPC components include by-products and phosphorylated intermediates of the formose cycle (Davis, 2006, 2013), characterized by a 1C driver (Breslow, 1959; Benner et al., 2010) and 2C sugar self-propagator (Fig. 7a), it is the apparent source of RPC 3C carbotide intermediates.

Figure 7b shows the RPC polypentotide is a ladder-like molecule with polyphosphate scaffolds cross-linked by cycle intermediates (Fig. 7b). Interspersing crosslinks and gaps produces a simple, transmissible, non-repeating binary sequence (Fig. 7c). Even this simple polycarbotide could, in principle, duplicate the function of a Tamura-Schimmel model RNA synthetase (Fig. 4b).

6.3 Compartmentalization, connectivity, and complexity

Back-tracking from bacteria and archaea populating the deepest branches in the tree of life, typified by *Aquifex* and *Methanopyrus* with genomes of only 1.55 to 1.69 million DNA bp, leads to a protocell as last common ancestor (LCA). Compartmentalization of the constituents of life within a bounded, microscopic, thermodynamically open, self-propagating antecedent undoubtedly occurred during prokaryote evolution. Not surprisingly, construction of a viable RNA-based protocell is currently a work in progress (Service, 2013). Archaea and bacteria membranes respectively contain phospholipids with an L- or D-phosphoglycerol scaffold and dual ether-linked, possibly branched, isoprenoid chains or acyl-linked unbranched fatty acid chains (Kates, 1977, Woese, 1987). This prompts the question: which came first? While archaea and bacteria undoubtedly shared a common ancestor before their membranes and a number of other features diverged, the possibility arises they diversified from a pre-membrane ancestor.

A degree of pre-membrane compartmentalization, with highly restricted spatial freedom, is exhibited by the cis-acting ribozymes that catalyze the reactions depicted in Fig. 6. Tracking-forward from the set of model replicators in Fig. 7, with spatial freedom restricted by cis-acting catalysis, leads to visualizing the LCA as a self-replicating polymer-complex. Ribosomes,

transamidosomes (Blaise et al., 2010), and the viral replication complex (Nagy and Pogany, 2012) individually form an integrated set of polymers that acts as macromolecular machine. Amino acid synthesis pathways during code formation provide additional evidence of functional polymer-complexes.

Takeover, on-block, of a four-step reaction sequence in the Val^5 path, occurred during formation of the Ile^7 synthesis pathway (Davis, 1999), and this indicates the Val^5 path incorporated a cassette of ribozymes, or synthesome particle (Davis, 2013a). Restriction of amino acid incorporation to the end-product, in a pathway with α-amino acid intermediates, as in the synthesis of Asp^1-derived amino acids, Thr^6, Ile^7, and Met^7 (dicarboxylated α-amino acid intermediates in Lys^{10} synthesis discounted), is also indicative of synthesome involvement. End-product incorporation also accounts for allocation of a codon 4-set, GCN, to Arg^9 at stage-III in code formation (restricted to doublet allocation on amino acid overprinting of pre-assigned codon 4-sets). It was apparently assigned initially to Orn^6, an α-amino acid intermediate, at stage-II (expansion through 4-set allocation). In principle, functional replisomes could defer membrane encapsulation, and protocell formation, until complexity levels were adequate for cell division. The membrane is seen here as one of a number of functional (non-replicating) structures produced by a complex polymer, whose monomer sequence encodes an algorithm for self-propagation, subject to the constraints of its own complexity and forces driving propagation.

Path invariants motivated the proposition that RPC components were polycarbotide monomers, in the pre-RNA era (Davis, 2012). Embedding them within a negatively charged polycarbotide conforms with the RPC initially occurring on a cationic mineral surface. Strand separation during poly-RPC replication (Fig. 7b) follows cleavage of the C2:C3 bond within the gem-diol form of 2-carboxy-3-keto-arabinose-1,5-bisphosphate (CKABP), in a retro-aldol type reaction (Safont et al., 2000). This is an extension of the mechanism responsible for spontaneous cleavage of the aldotetrose C2:C3 bond (Figs. 7a), to produce two GO molecules, in the aqueous phase formose cycle (Breslow, 1959). Thereafter, double strand molecules with counter-oriented RuP, or RP, monomers (Figs. 7d,e), respectively stabilized by sets of (O2••H-O3, O3-H••O2) or (O2-H••O4,

O4••H-O2)[4] H bonds (Ulyanov and James 2010), evolved in this scenario. They can be anticipated, from the physicochemical principles governing evolution (Davis. 1996, 1998, 2000), to have enhanced polycarbotide replicative and functional (structural, catalytic) properties.

This depicts replication as the product of evolution by accretion, where the replicator at one stage became the scaffold for its successor, in a 'piggy-back'-like mechanism:

formose cycle: [aldotetrose, C2:C3] (Fig. 7a) ⟶

⟶ RPC ladder: [RuP, C2:C3] polyphosphate scaffold, sugar/gap sequence (Fig. 7b,c)

⟶ poly(RuP) flat duplex: [RuP, O3-H••O2] poly(RuP) scaffold, L- RuP/R-RuP sequence (Fig. 7d)

⟶ poly(RP) flat duplex: [RP, O2-H••O4] scaffold 3-5 edge, 2-4/4-2 RP-edge sequence (Fig. 7e)

⟶ RCC ladder: [Cit, C2:C3] poly(RP) scaffold, intermediates sequence (Fig.7f);

⟶ RNA coiled duplex: [A:U, G:C, N6-H••O4 N1••H-N3, O6••H-N4 N1-H••N3 N2-H••O2] poly(RP) scaffold, A,G,C,U sequence (Fig 7g).

configurations listed refer to replicative-form; within brackets are molecule(s) and its bond severed during strand separation.

Trans-acting RNA ligase with processivity, tC9Y, compartmentalized in a eutectic mixture of ice and water and tethered to a primer (Attwater et al. 2013), replicated an RNA strand longer than itself (202 nt). A Group I intron, on the other hand, utilizes a cis-acting ligase (Hougland et al. 2006). When taken together, early self-replication by RNA appears possible, giving added credence to the RNA World scenario (Gilbert, 1985). Since carbotide transients arise in the RNA ligase reaction (Fig. 6a), carbozyme catalyzed polycarbotide self-replication plausibly preceded ribozyme catalyzed RNA replication. Surface-bound, cis-acting polymeric catalysts would, in principle, avoid the randomness that de Duve and Miller (1991) saw as a kinetic impediment to the surface reaction system envisioned by Wächtershäuser (1992). They also noted that favorable chain extension thermodynamics in a surface system elevated the risk of entrapment by irreversible adhesion to the surface. However, a polycarbotide replicator that self-annealed into a series of hairpin loops,

[4] Equivalent to strong pyranose O3-H••O5 (intra-chain) H bonds in cellulose (Poletto et al. 2013).

familiar from the duplex-avoidance strategy exhibited by competitively replicated variants in the Qβ replicase system (Davis 1995a, b, 1998), could likewise avoid surface entrapment. In addition to circumventing the de Duve-Miller restrictions, polymer formation provided a means to stabilize labile metabolites (Robinson et al., 1973, Davis, 2006), a significant feature at the origin of life under early Earth conditions (Yakhnin, 2013)

Six inorganic C-fixing pathways are currently known for biomass formation in anaerobic prokaryotes (Fuchs, 2011): RPC, RCC, dicarboxylate/4-hydroxybutyrate cycle (DCHB), 3-hydroxypropionate/4-hydroxybutyrate cycle (HPHB), 3-hydroxypropionate bi-cycle (HPBC), and reductive acetyl-coenzyme A (RACA) pathway. Five of them, listed below as net reactions, are autocatalytic, with intermediates containing an invariant terminal phosphate or carboxyl,

RPC: $CO_2 + 2ATP + 2H_2O + NADPH + H^+ + DHAP + GAP + [\mathbf{PGA}] \xrightarrow{[-PO_4]} 2\,[\mathbf{PGA}] + EP + NADP^+ + 2ADP + 2PO_4H^{2-}$

RCC: $3CO_2 + HCO_3^- + 3ATP + 3H_2O + 2NADPH + NADH + 3H^+ + 4Fd_{red} + [\mathbf{OA}] \xrightarrow{[-COOH]} 2[\mathbf{OA}] + 4Fd_{ox} + 2NADP^+ + NAD^+ + 2ADP + AMP + 4PO_4H^{2-}$

DCHB: $CO_2 + HCO_3^- + 3ATP + 5H_2O + 3NADPH + 2NADH + 5H^+ + 2Fd_{red} + CoA + [\mathbf{CH_3COS}\text{-CoA}] \xrightarrow{[-COOH]} 2[\mathbf{CH_3COS}\text{-CoA}] + 2Fd_{ox} + 3NADP^+ + 2NAD^+ + ADP + 2AMP + 5PO_4H^{2-}$ (1)

HPHB: $2HCO_3^- + 4ATP + 5H_2O + 5NADPH + NADH + 6H^+ + CoA + [\mathbf{CH_3COS}\text{-CoA}] \xrightarrow{[-COOH]} 2[\mathbf{CH_3COS}\text{-CoA}] + 5NADP^+ + NAD^+ + 2ADP + 2AMP + 6PO_4H^{2-}$

HPBC: $2HCO_3^- + 3ATP + 4H_2O + 3NADPH + 3H^+ + UQ_{ox} + CH_3CH_2COS\text{-}CoA + [\mathbf{CH_3COS}\text{-CoA}] \xrightarrow{[-COOH]} 2[\mathbf{CH_3COS}\text{-CoA}] + CH_3C{=}OCOOH + 3NADP^+ + UQ_{red} + 2ADP + AMP + 4PO_4H^{2-}$

RACA: $ATP + CO_2 + NADPH + NADH + 2H^+ + 2Fd_{red} + THF \longrightarrow CH_3\text{-}THF + ADP + PO_4H^{2-} + 2Fd_{ox}$

$CH_3\text{-}Co(III)FeSP \quad Co(I)FeSP$

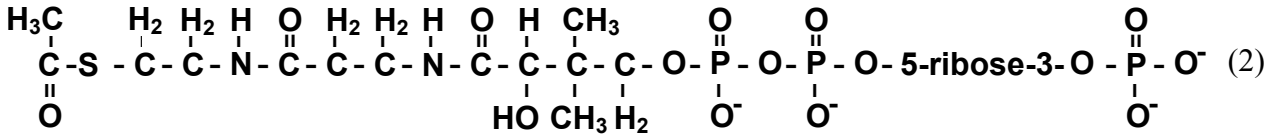
$CO_2 + 2Fd_{red} + 2H^+ + CoA \longrightarrow CH_3OS\text{-}CoA + H_2O + 2Fd_{ox}$

bold letters highlight an autocatalytic component or invariant group, and brackets signify an inferred polymer scaffold. The sixth inorganic C fixing pathway, RACA, contains two non-autocatalytic

cycles, for recovery of tetrahydrofolate (THF), in bacteria[5], and Co^{1+}-corrinoid-iron-sulfur protein (Co(I)FeSP). A linear reaction sequence, involving two C1 molecules, produces acetyl-CoA.

Unlike the foregoing C-fixation cycles, the RACA pathway contributes either energy or biomass (Bar-Even et al., 2011). The former through conversion to acetyl-phosphate, recovering an ATP hydrolyzed in attaching formate to THF and indirectly through coupling an oxidation-reduction step, involving a THF derivative, to a membrane ATPase sodium pump. And, the latter follows from pyruvate formation. Formation of the RACA path early in the emergence of life (Martin and Russell, 2007; Fuchs, 2011; Nitschke and Russell, 2013; Herschy et al. 2014; Sousa and Martin, 2014) is supported by: (i) its dual energy/biomass functionality; (ii) direct assimilation of reduced C1 substances (formate, CO); (iii) reliance on metalloproteins (carbon monoxide dehydrogenase, acetyl-CoA synthase, ferredoxin, Co^{1+} or Co^{3+} bearing corrinoid iron sulfur protein); (iv) participation in the autocatalytic generation of activated acetate by the DCHB, HPHB, and HPBC (Eqn. 1) and oxaloacetate by the RCC, together with (v) its role in the synthesis of fatty acids, sterols, and sugar, via pyruvate, among anaerobic chemolithotrophic prokaryotes.

Reductive acetyl-CoA pathway antiquity, evident from its extensive connectivity, presupposes a pre-RNA antecedent. With acetate and adenine at the thioester and ribotide ends of CoA, respectively, nucleobase participation in any transfer reaction can be reasonably discounted. Deadenosine-acetyl-coenzyme A (acetyl-β-mercaptoethylamine-pantothenate-5-diphospho-ribosyl-3-phosphate) is thus a likely antecedent;



A phosphate at ribose-C3 in this molecule points to prior attachment to a polyribotide (§3). Earlier CoA antecedents, based on comparative genomics (Genschel, 2004), include 4'-phospho-N-

[5] Tetrahydromethanopterin (THMP) in archaea.

pantothenoylcysteine, that likewise bear a terminal phosphate suitable for a polyribotide scaffold. With amino acids L-Cys, L-Val, and L-Asp derived β-Ala as precursors and an invariant attributed to a former scaffold, synthesis of CoA and its antecedents broadly resembles the nascent nucleobase pathway (§3). Use of an L-enantiomer of each amino acid precursor may be viewed as evidence that the nascent coenzyme pathway arose after a mechanism for homochirality had evolved. By forming the scaffold for synthesis of CoA and nuclebases, ribose evidently predated both. As ribose is a by-product of the RPC, this cycle appears to be of great antiquity. In accord with this it is uniquely independent of CoA among C fixing pathways (Eqn. 1).

Intermediates in CoA dependent C-fixing pathways, RCC, DCHB, HPHB, and HPBC, differ from the RPC in having an invariant terminal carboxyl in place of phosphate. A carboxyl invariant also conforms, however, with the former anchoring of a polycarbotide;.bonding being through the scaffold hydroxyl, as in acylated tRNA. Invariant carboxyls at each terminus of RCC intermediates base the early cycle on a ladder-like replicator cross-linked by intermediates (Fig. 7f), analogous to the RPC. Direct and indirect (via metylmalyl-CoA) bicycle HPBC paths, from succinyl-CoA to acetyl-CoA, contain a similar pattern of terminal carboxyl, or carboxyl-CoA groups. More generally, each of the CoA dependent C-fixing cycles oscillates between acetyl-CoA and succinyl-CoA (Bar-Even et al. 2011), revealing them to be related. The number of distinct C-fixing metabolic cycles thus reduces to two: CoA dependent or independent. A subdivision in these cycles that can be most simply attributed to the former (RCC, DCHB, HPHB, HPBC) having evolved after CoA, while the latter (RPC) appeared some time before the coenzyme. The antiquity of the RPC, RCC, and RACA (THF form) is apparent from their occurrence in chemolithotropic bacteria (Bar-Even et al. 2011), with the RPC accounting for most C-fixation, among all species.

Enzymes catalyzing THF/THMP dependent reduction of CO_2 to a methyl group differ between acetogenic bacteria and methanogenic archaea (Madden, 2000; Sousa and Martin, 2014). Evidently this part of the RACA path (Eqn. 1) arose independently, through convergent evolution, in each prokaryote domain, after they had diverged from the LCA. Homology between carbon monoxide dehydrogenase (CODH) and acetyl-CoA synthase (ACS), by contrast, implies acetyl-CoA

synthesis preceded prokaryote divergence. This dissociates evolution of the nascent RACA path from chemiosmotic energy generation and reduces it to a dependence on substrate level energy. It also leaves the initial source of the methyl group in acetyl-CoA synthesis unspecified, although geological sources (Etiope et al., 2011) would provide a plausible source. Even if the denitrifying methanotropy reaction sequence proposed by Nitschke and Russell (2013) preceded prokaryote divergence from the LCA, it would not place this non-autocatalytic pathway at the origin of metabolism.

Activated thioacetic acid (CH_3-CO-SCH_3) can form from carbon monoxide and methane thiol (CH_3-SH) under simulated hydrothermal conditions (!00^oC, 7 days), in the presence of a coprecipitated FeS and NiS slurry (Huber and Wächtershäuser, 1997). Occurrrence of FeS and NiS clusters in the active site of CODH and ACS invites the conjecture that a standard reaction, or reaction sequence, of this kind initiated the emergence of life. The residue sequence of these enzymes can be viewed as a folding algorithm that enables each active site to possess the required substrate specificity and metal ligands. That CODH and ACS should mimic a reaction present under abiotic conditions is a familiar outcome of evolution (Wickler, 1968). In the present context, it implies that early replicating entities detected the forces driving the abiotic reaction and evolved to utilize them to drive propagation. In this sense, the biocatalyst (carbozyme, ribozyme, enzyme) does not create a new reaction, but optimizes an existing one (Eschenmoser and Loewenthal, 1992). This is achieved, however, in the context of generating sequence complexity sufficient to catalyze the reaction. Carbotide era sources of energy and biomass, by comparison, are the ubiquitous phosphodiester bond and sugar molecules (derived from CO_2), respectively, as in extant metabolism.

As a potential source of energy and matter for initiating life, the activated acetate thioester bond in acetyl-CoA and its 1C source, two CO_2 molecules, apply, respectively. As acetyl-CoA synthesis involves a standard reaction sequence (Eqn. 1), it lacks the self-organizing capacity to generate the threshold molecular complexity required for primal life. A 170 residue sequence (segments G260 – V338 and I 437 – V527, *Carboxydothermus hydrogenoformans* numbering) in cluster C spans the

active site of CODH II (Dobbek et al., 2001). It has a *prima facie* sequence complexity of 170 vigesimal units (complexity equivalent to 20^{170} possible sequences); basing the complexity on amino acid frequencies within the sequence yields a marginally lower value, 158.2 vigesimal units, where

$$\mathcal{C} = -\mathrm{N} \sum_{i=1}^{20} p_i \log_{20} p_i \qquad 0 \le p_i \le 1 \tag{3}$$

$\mathcal{C}$ denotes total sequence complexity, N is sequence length, and *p*i is the frequency of amino acid of type i. In addition to its intrinsic complexity, the residue sequence exhibits homochirality. It is apparent that a credible theory on the origin of life should incorporate a complexity generating mechanism, together with acceptable sources of energy and matter. With respect to the role of acetyl-CoA synthesis in the origin of life, as the product of a standard, non-autocatalytic reaction, its invariant carbotide (Eqn. 2) suggests the coenzyme initially arose on a branch path within a pre-RNA self-propagating polycarbotide complex.

6.4 Source of homochirality

Chiral asymmetry manifest *in vivo* by the utilization of D-sugars and L-amino acids had its origin in self-propagation, according to Frank (1953), and he outlined a scheme to establish the proposed connection. Comparatively recent experiments involving 5-pyrimidyl alkonol synthesis in the presence of a chiral selective catalyst (Soai et al., 1995; Blackmond et al. 2004) succeeded in substantiating chiral symmetry breaking through autocatalytic amplification of an enantiomeric excess. A physically induced autocatalysis in sodium chlorate crystallization also resulted in chiral asymmetry (Kondepudi et al, 1990). Interestingly, amino acid L-Ala catalyzed homochiral synthesis of sugar D-threose from the autocatalyst glycoaldehyde (Pizzarello and Weber, 2004). The Tamura-Schimmel ribozyme synthetase (§4), with a D-ribotide scaffold, was additionally noted to preferentially catalyze cofactor acylation with the L-enantiomer of Ala, for steric reasons.

The replicative form of poly(triose-phosphate), depicted as CKABP monomers within poly(RuBP)

in Fig. 7b,c, was previously seen as a chiral filter (Davis, 2012, 2013b). This derived from interference by the CKABP carboxylate, which prevented contact between terminal phosphates in pairs of mixed chiralty, precluding phosphate bond formation and, consequently, extension of the polyphosphate scaffold. Figure 8 shows that formation of the polyphosphate scaffolds in poly(RuBP) requires RuBP homochirality. Interference by hydroxyls at C3 in RuBP pairs of mixed

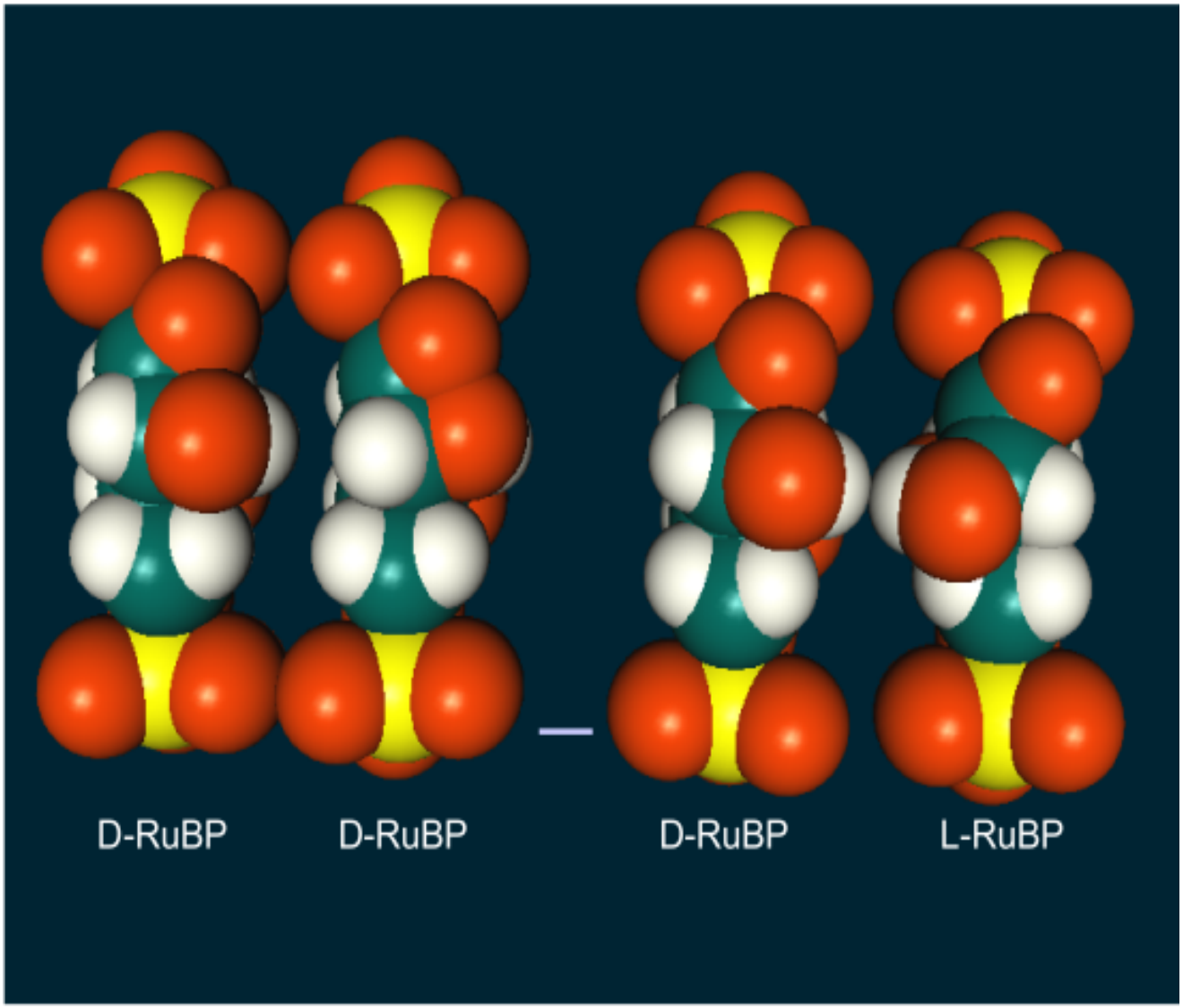


Fig. 8. A CPK model showing steric hindrance between D and L RuBP enantiomers. In contrast to the ketopentol pair with same chirality, OH groups at C3 in the mixed chirality pair prevent contact between terminal phosphate groups, blocking.growth of the polyphosphate scaffold at each end. The RuBP models were constructed as in Fig. 4. Yellow, P atom; green, C; orange, O; white, H. Bar length, 1 Å.

chirality prevent condensation of adjoining phosphates, thereby blocking growth of the polyphosphate scaffold. Figure 8 demonstrates that formation of the polyphosphate scaffolds of poly(RuBP) requires RuBP homochirality. Application of the model for pre-RNA metabolism, based on the distribution of invariants in the pathways of central metabolism, has thus identified a long sought after source of *in vivo* sugar homochirality. The link established between synthesis of a D-sugar in the presence of an L-amino acid (Pizzarello and Weber, 2004), additionally raises the possibility that the converse relationship, evident in cofactor acylation by the Tamura-Schimmel ribozyme synthetase (§4), accounts for amino acid homochirality, *in vivo*.

6.5 Emergence of life

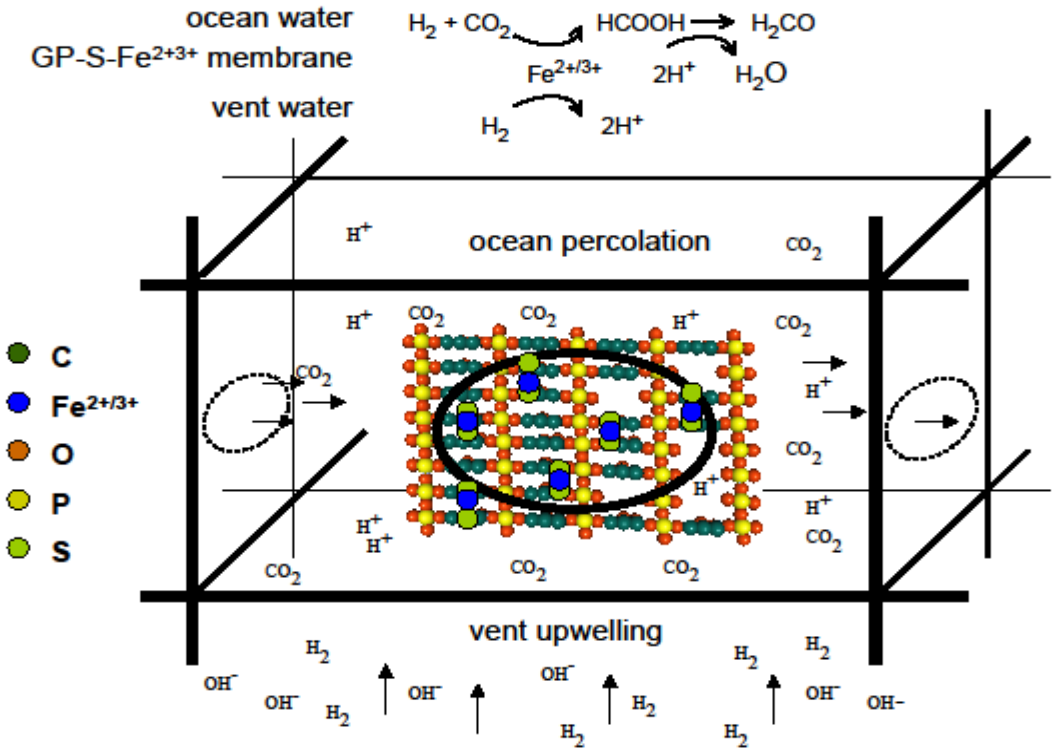
The emergence of life on the early Earth occurred at a time that is now far distant from real time; around four billion years distant. In addition, for all its breath-taking diversity, the essential universality of the genetic code and, even earlier, reaction sequences of central metabolism, establish that the spontaneous initiation of life is a rare event, possibly even a singular event in the long history of this planet. For these reasons, it may never be possible to uncover preserved traces of the first life form. Yet, we know life appeared comparatively rapidly after the Earth formed. Furthermore, it is subordinate to the laws of physics and chemistry. And, considering the physicochemical simplicity of the processes central to life, the prospects of recreating life in the laboratory (Alamada and Szostak, 2013; Herschy et al, 2014) appear good.

Interpreting invariants in the pathways of central metabolism as vestiges of an early cofactor/ scaffold attachment site, based on features observed in reconstructed tRNA-dependent amino acid synthesis pathways (Davis, 2013a), has led to the model for the evolution of replication proposed here. As long anticipated, it places the spontaneous, autocatalytic synthesis of the 2C sugar, glycoaldehyde, from a 1C source, at its root. From there, the first 3C carbotide replicator evolved and led to the 5C ribotide replicator that became the RNA scaffold. This identifies the experimentally demonstrable self-organizing capacity of autocatalytic processes as an invariant of all life. It thus marks the origin of life as the point in time, when the first polymer encoded a linear

algorithm for a structure and function that promoted its own replication. By this criterion, the RPC replicator (Fig. 7b,c) is the first life-form. A model of the evolution of replication thus respresents a model of the evolution of evolution.

When approached from a geochemical perspective, free energy in a proton gradient at the interface between an alkaline hydrothermal vent upwelling and acidified Hadean ocean appears likely to have helped propel the origin of life (Russell et al., 1993). Most energy for metabolism, in all cells, is derived from a proton gradient at the cell membrane, used to generate high-energy phosphate bonds (Lane et al., 2010). Since hydrolysis of ATP phosphodiester bonds is the contingent source of energy in most contemporary biochemical reactions, utilization of a remote, non-substrate-level, membrane proton gradient can be viewed as a subsequent development, notwithstanding its significance as an ATP source. Various amphipathic lipids are capable of self-assembly in an aqueous medium; however, they are incapable of self-organization beyond forming membrane vesicles. Being non-linear, membranes cannot replicate. And, being non-aqueous phase fluids, they retain no information to transmit. Membrane embedded proteins, including the lipoprotein subunit of $[H^+]$-ATP synthetase, contain information and their genes replicate. But their presence at the origin of life would be achronistic. The residue profile of pre-LCA lipoprotein, for example, matches that of a stage-7 code, and the enzyme aqueous phase subunits arose later (Davis, 2002).

Membrane phospholipids incorporate a glycerol-3-P (GP) scaffold. When crosslinked, poly(GP) forms a membrane of variable porosity (Davis, 2013b). In this circumstance, membrane and 3C sugar autocatalyst, GAP, are directly linked. Substituting a crosslinked thiolated-GP membrane, with coordinately bonded $Fe^{2+/3+}$ atoms, for the mineral FeS membrane envisioned by Russell et al. (1993) at an alkaline hydrothermal vent, in principle, could yield a primal dehydrogenase that is directly linked to the self-organizing capacity of triose-P autocatalyst (Fig. 9).



. **Fig. 9.** A primal FeS 'dehydrogenase' coupled to a triose-P autocatalyst. A semi-permeable crosslinked poly(GP) membrane at an alkaline hydrothermal vent covers a 'chamber' pore in a vent rock. Free energy contributed by the proton gradient across the membrane partitioning an acidic Hadean ocean and vent upwelling drives the synthesis of formaldehyde from CO_2. The proposed mechanism couples a geophysical free energy source to self-organizing capacity of GP, derived from related autocatalyst GAP, produced by RPC replication (Fig. 7b).

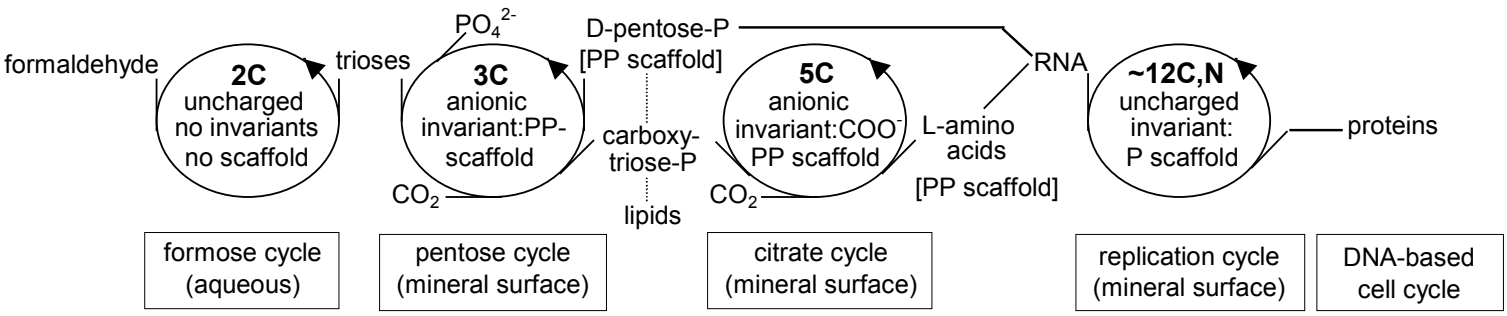


Fig. 10. Stages in the proposed evolution of replication based autocatalytic biomolecular cycles from the spontaneous formose cycle to the cell cycle. Each successive stage builds upon its antecedent.

In addition to connecting a principal geochemical free energy source on the early Earth with the self-organizing capacity of pre-RNA replicators, the proposed replication of polycarbotides integrates the biomolecular cycles depicted in Fig. 10 and links them to the spontaneous autocatalytic, sugar-forming formose cycle.

Acknowledgement

Professor Martin Rees motivated my interest in the origin of life, during a conversation at Carlton House Terrace; Professor Hans Christensen and associates, at Denmark Technical University, skillfully validated a prediction of mine on an early protein structure; and advice was received from Dr Michael Russell at various stages in my research.

References

Agmon I, Bashan A, Zarivach R, Yonath A. 2005 Symmetry at the active site of the ribosome: Structural and functional implications. *Biol. Chem.* **386**, 833-844 (doi: http://dx.doi.org/10.1515/BC.2005.098)

Agmon I, Davidovich C, Bashan A, Yonath A. 2009 Identification of the prebiotic translation apparatus within the contemporary ribosome. http://precedings.nature.com/documents/2921/

Adamala K, Szostak JW. 2013 Non-enzymatic template directed RNA synthesis inside model protocells. *Science* **342**, 1098-1011. (doi: http://dx.doi.org/10.1126/science.1241888)

Attwater J, Wochner A, Holliger P. 2013 In-ice evolution of RNA polymerase ribozyme activity. *Nature Chem.* **5**, 1011-1018. (doi: http://dx.doi.org/10.1038/nchem.1781)

Banas P, Walter NG, Sponer J, Otyepka M. 2010 Protonation states of the key active site residues and structural dynamics of the *glmS* riboswitch as revealed by molecular dynamics. *J. Phys Chem*. B **114**, 8701-8712. (doi: http://dx.doi.org/10.1021/jp9109699)

Bar-Even A, Noor E, Milo R. 2011 A survey of carbon fixation pathways through a quanitative lens. *J. Exp. Botany* 1-18. (doi: http://dx.doi.org/10.1093/jxb/err417)

Benner SA, Kim H-J, Kim M-J, Ricardo A. 2010 Planetary organic chemistry and the origins of biomolecules. *Cold Spring Harbor Perspectives in Biology* **2(7)**: a003467.

Bernhardt HS, Sandwick RK. 2014 Purine biosynthetic intermediate-containing ribose-phosphate polymers as evolutionary precursors to RNA. *J. Mol. Evol.* **79**, 91-104 (doi: http://dx.doi.org/10.1007/s00239-014-9640-1)

Blackmond D G. 2004 Asymmetric autocatalysis and its implications for the origin of homochirality. *Proc. Natl. Acad. Sci. USA* **101**, 5732-5736 (doi: http://dx.doi.org/10.1073/pnas.0308363101)

Blaise M, Bailly M, Frechin M, Behrens MA, Fischer F, Oliviera CLP, Becker HD, Pederson JS, Thirup S, Kern D. 2010 Crystal structure of a transfer-ribonucleoprotein particle that promotes asparagine formation *EMBO J*, **29**, 3118-3129. (doi: http://dx.doi.org/10.1038/emboj.2010.192)

Breslow R. 1959 On the mechanism of the formose reaction. *Tetrahedron Lett.* **21**, 22-26. (doi: http://dx.doi.org/10.1016/S0040-4039(01)99487-0)

Breslow R 2011 A likely possible origin of homochirality in amino acids and sugars on prebiotic earth. *Tetrahedron Lett.* **52**, 2028-2032. (doi: http://dx.doi.org/10.1016/j.tetlet.2010.08.094)

Butelrow A. 1861 Bildung einer zuckerartigen Substanz durch Synthese. *Leibigs Ann. Chem.* **120**, 295-298. (doi: http://dx.doi.org/10.1002/jlac.18611200308)

Brooks DJ, Fresco JR, Lesk AM, Singh M. 2002 Evolution of amino acid frequencies in proteins over deep time: Inferred order of introduction of amino acids into the genetic code. *Mol. Biol. Evol*. **19**, 1645-1655. (doi: http://dx.doi.org/10.1093/oxfordjournals.molbev.a003988)

Calvin M. 1969 *Chemical Evolution: Molecular Evolution towards the Origin of Living Systems on the Earth and Elsewhere* London: Clarendon Press.

Cochrane JC, Lipchock JC, Strobel SA. 2007 Structural investigation of the *GlmS* ribozyme bound to its catalytic cofactor. *Chem. Biol.* **14**, 97-105. (doi: http://dx.doi.org/10.1016/j.chembiol.2006.12.005)

Davidovich C, Belousoff M, Bashan A, Yonath A. 2009 The evolving ribosome: from non-coded peptide bond formation to sophisticated translation machinery. *Res. Micribiol.* **160**, 487-492. (doi: http://dx.doi.org/10.1016/j.resmic.2009.07.004)

Davis BK. 1978a Rate of polymer formation and entropy production during competitive replication. *J. Mol. Evol.* **10**, 325-338. (doi: http://dx.doi.org/10.1007/BF01734222)

Davis BK. 1978b Change in polymerization rate and entropy production during competitive replication. In *Biomolecular Structure and Function* Eds. Agris PF, Loeppky RN, Sykes BD New York: Academic Press pp. 607-610.

Davis BK. 1995a Stability of replicative form and fitness among RNA variants transcribed by Qβ replicase. *Proc. R. Soc. Lond.* **B 260**, 39-43. (doi: http://dx.doi.org/10.1098/rspb.1995.0056)

Davis BK 1995b Significance of strand configuration in self-replicating RNA molecules. *Phil. Tans R Soc.Lond.* **B 350**, 345-352. (doi: http://dx.doi.org/10.1098/rstb.1995.0169)

Davis BK. 1996 A theory of evolution that includes prebiotic self-organization and episodic species formation. *Bull. Math. Biol.* **58**, 65-97 (doi: http://dx.doi.org/10.1007/BF02458283)

Davis BK. 1998 The forces driving molecular evolution. *Prog. Biophys Mol. Biol*. **69**, 83-150. (doi: http://dx.doi.org/10.1016/S0079-6107(97)00034-5).

Erratum. *Prog. Biophys. Mol. Biol.* **71**, I-II. (doi: http://dx.doi.org/10.1016/S0079-6107(98)00053-4)

Davis BK. 1999 Evolution of the genetic code. *Prog. Biophys. Mol. Biol.* **72**, 157-243 (doi: http://dx.doi.org/10.1016/S0079-6107(99)00006-1)

Davis BK. 2000 Darwinian aspects of molecular evolution at sublinear propagation rates. *Bull. Math. Biol.* **62**, 675-694. (doi: http://dx.doi.org/10.1006/bulm.2000.0204)

Erratum *Bull. Math. Biol.* **62**, 1001. (doi: http://dx.doi.org/10.1006/bulm.2000.0173)

Davis BK. 2002 Molecular evolution before the origin of species. *Prog. Biophys. Mol. Biol.* **79**, 77-132. (doi: http://dx.doi.org/10.1016/S0079-6107(02)00012-3)

Davis BK. 2006 Making sense of the genetic code with the path-distance model. In *Leading-Edge Messenger RNA Research Communications* Ed. Ostrovskiy MH New York: Nova Science pp 1-32

Davis BK. 2008 Comments on the search for the source of the genetic code. In *Messenger RNA Research Perspectives* Ed. Takeyama T New York: Nova Science pp. 1-8

Davis BK. 2008 Imprint of early tRNA diversification on the genetic code: Domains of contiguous codons read by related adaptors for sibling amino acids In *Messenger RNA Research Perspectives* Ed. Takeyama T New York: Nova Science pp. 35-79

Davis BK. 2009 On mapping the genetic code. *J. Theoret. Biol.* **259**, 860-862. (doi: http://dx.doi.org/10.1016/j.tbi.2009.05.009)

Davis BK. 2011 Genetic code domains conserve the imprint of tRNA cofactors encoded to direct the synthesis of encoded amino acids. url: http://www.archive.org/details/GeneticCodeDomains

Davis BK. 2012 Replicative-form of poly(triose-phosphate). url: http://www.archive.org/details/Replicative-formOfPolytriose-phosphate

Davis BK. 2013a Making sense of the genetic code with the path-distance model based on tRNA-dependent pathways. url: http://www.archive.org/details/MakingSenseOfGeneticCode

Davis BK. 2013b An organizing principle for the emergence of life. url: http://www.archive.org/details/OrganizingPrinciple

Decker P, Speidel A. 1972 Open systems which can mutate between several steady states ("biods") and a possible prebiological role of the autocatalytic condensation of formaldehyde. *Z. Naturforsch.* **27b**, 257-263.

de Duve C, Miller SL. 1991 Two-dimensional life. *Proc. Natl. Acad. Sci. USA* **88**, 10014-10017. (doi: http://dx.doi.org/10.1073/pnas.88.22.10014)

Diederichsen U. 1996 Pairing properties of alanyl peptide nucleic acids containing an amino acid backbone with alternating configuration. *Angew. Chem. Int. Ed. Engl.* **35**, 445-448. (doi: http://dx.doi.org/10.1002/anie.199604451)

Dobbek H, Svetlitchnyi V, Gremer L, Huber R, Meyer O. 2001 Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster. *Science* **293**, 1281-1285. (doi: http://dx.doi.org/10.1126/science.1061500)

Doctor VM, Oro J. 1972 Non-enzymatic β-decarboxylation of aspartic acid. *J. Mol. Evol.* **1**, 326-333. (doi: http://dx.doi.org/10.1007/BF01653961)

Egholm M, Buchardt O, Nielsen PE, Berg RN. 1992 Oligonucletide analogues with an achiral peptide backbone. *J. Am. Chem. Soc*. **114**, 1845-1897. (doi: http://dx.doi.org/10.1021/ja00031a062)

Eigen M. 1971 Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465-523. (doi: http://dx.doi.org/10.1007/BF00623322)

Eiler S, Dock-Bregeon A-C, Moulinier L, Thierry J-C, Moras D. 1999 Synthesis of aspartyl-tRNA^{Asp} In Escherichia coli – a snapshot of the second step. *EMBOJ* **18**, 6532-6541. (doi: http://dx.doi.org/10.1093/emboj/18.22.6532)

Eriani G, Delarue M, Poch O, Gangloff J, Moras D. 1990 Patrition of aminoacyl-tRNA synthetases into two classes on the basis of two mutually exclusive sets of sequence motifs. *Nature* **347**, 203-206. (doi: http://dx.doi.org/10.1038/347203a0)

Eshenmoser A, Loewenthal E. 1992 Chemistry of potentially prebiological natural-products. *Chem. Soc. Rev.* **21**, 1-16. (doi: http://dx.doi.org/10.1039/cs9922100001)

Etiope G, Schoell M, Hosgormez H. 2011 Abiotic methane flux from the Chimaera seep and Tekirova ophiolites (Turkey): understanding gas exhalation from low temperature serpentinization and implications for Mars. *Earth Planet. Sci. Lett.* **310**, 96-104. (doi: http://dx.doi.org/10.1016/j.epsl.2011.08.001)

Fisher RA. 1930 *The Genetical Theory of Natural Selection* 2nd edn. New York: Dover.

Forconi M, Piccirilli JA, Herschlag D. 2007 Modulation of individual steps in group I intron catalysis by a peripheral metal ion. *RNA* **13**, 1656-1667. (doi: http://dx.doi.org/10.1261/rna.632007)

Frank CF. 1953 On spontaneous asymmetric synthesis. *Biochim. Biophys. Acta* **11**, 459-463. (doi: http://dx.doi.org/10.1016/0006-3002(53)90082-1)

Fuchs G. 2011 Alternative pathways of carbon dioxide fixation: Insights into the early evolution of

life? *Annu. Rev. Microbiol.* **65**, 631-658. (doi: http://dx.doi.org/10.1146/annurev-micro-090110-102801)

Genschel, U. 2004 Coenzyme A biosynthesis: Reconstruction of the pathway in archaea and an evolutionary scenario based on comparative genomics. *Mol. Biol. Evol.* **21**, 1242-1251. (doi: http://dx.doi.org/10.1093/molbev/msh119)

Gilbert W 1986 Origin of Life: The RNA World. *Nature* **319**. 618. (doi: http://dx.doi.org/10.1038/319618a0)

Gindulyte A, Bashan A, Agmon I, Massa L, Yonath A, Karle J. 2007 The transition state for formation of the peptide bond in the ribosome *Proc. Natl. Acad. Sci USA* **103**, 13327-13332 (doi: http://dx.doi.org/10.1073/pnas.0606027103)

Greenberg DM. 1969 *Metabolic Pathways* Ed. Greenberg DM Vol. 3 New York: Academic Press, pp. 237-315.

Groebke K, Hunziker J, Fraser W, Peng L, Diederichsen U, Zimmermann K, Holzner A, Leuman C, Eschenmoser A. 1998 Why pentose- and not hexose-nucleic acids? Part V Purine-purine pairing in homo-DNA: Guanine, isoguanine, 2,6-diaminopurine, and xanthine. *Helv. Chim. Acta* **81**, 375-474. (doi: http://dx.doi.org/10.1002/hlca.19980810302)

Haugen P, Simon DM, Bhattacharya D, 2005 The natural history of group I introns. *Trends in Genetics* **21**, 111-119 (doi: http://dx.doi.org/10.1016/j.tig.2004.12.007)

Herschy B, Whicher A, Camprubi E, Watson C, Datnell L, Ward J, Evans RG, Lane N. 2014 An origin-of life reactor to simulate alkaline hydrothermal vents. *J. Mol. Evol*. **79**, 213-227. (doi: http://dx.doi.org/10.1007/s00239-014-9658-4)

Hinshelwood CN. 1946 *The Chemical Kinetics of the Bacterial Cell* London: Clarendon Press.

Holm NG, Neubeck A. 2009 Reduction of nitrogen compounds in oceanic basement and its implications for HCN formation and abiotic organic synthesis. *Geochemical Transactions* **10:9**

Honma H. 1996 High ammonium contents in the 3800 Ma Isua supracrustal rocks, central West Greenland. *Geochim. Cosmochim. Acta* **60**, 2173-2178. (doi: http://dx.doi.org/10.1016/0016-7037(96)00083-X)

Hougland JL, Piccirilli JA, Forconi M, Lee J, Herschlag D. 2006 *How the group I intron works: A case study of RNA Structure and Function* Cold Spring Harbor: Cold Spring Harbor Monograph Series **43**, 1-133

Huber C, Wächtershäuser G. 1997 Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* **276**, 245-247. (doi: http://dx.doi.org/10.1126/science.276.5310.245)

Kates M 1977 The phytanyl ether-linked polar lipids and isoprenoid neutral lipids of extremely halophilic bacteria. *Prog. Chem. Fats Other Lipids* **15**, 301-342. (doi: http://dx.doi.org/10.1016/0079-6832(77)90011-8)

Kelley DS, Karson JA, Blackman DK, FruÈh-Green GL, Butterfield DA, Lilley MD, Olson EJ, Schrenk MD, Roe KK, Lebon GT, Rivizzigno P. 2001 An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30N. *Nature* **412**, 145-149. (doi: http://dx.doi.org/10.1038/35084000)

Klein DJ, Wilkinson SR, Been MD, Fette-D'Amare AR 2007 Requirement of helix P2.2 and nucleotide G1 for positioning the cleavage site and cofactor of the *glmS* ribozyme. *J. Mol. Biol*. **373**, 178-189. (doi: http://dx.doi.org/10.1016/j.jmb.2007.07.062)

Kondepudi DK, Kaufman R, Singh N. 1990 Chiral symmetry breaking in sodium chlorate crystallization. *Science* **250**, 975-976. (doi: http://dx.doi.org/10.1126/science.250.4983.975)

Kramer FR, Mills DR, Cole PE, Nishara T, Spiegelman S. 1974 Evolution *in vitro*: sequence and phenotype of a mutant RNA resistant to ethidium bromide. *J. Mol. Biol.* **89**, 719-738. (doi: http://dx.doi.org/10.1016/0022-2836(74)90047-3)

Kvenvolden K, Lawless J, Pering K, Peterson E, Flores J, Ponnampurema C, Moore C. 1970 Evidence for extraterrestrial amino acids and hydrocarbons in the Murchison meteorite. *Nature* **228**, 923-926. (doi: http://dx.doi.org/10.1038/228923a0)

Lane N, Allen JF, Martin W. 2010 How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays* **32**, 271-280. (doi: http://dx.doi.org/10.1002/bies.200900131)

Lang SQ, Butterfield DA, Schulte M, Kelley DS, Lilley MD. 2010 Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim. Cosmochim. Acta* **74**, 941-952. (doi: 10.1016/j.gca.2009.10.045)

Leontis NB, Westhof E. 2001 Geometric nomenclature and classification of RNA base pairs. *RNA* **7**, 499-512. (doi: http://dx.doi.org/10.1017/S1355838201002515)

Martin W, Russell MJ. 2007 On the origin of biochemistry at an alkaline hydrothermal vent. *Phil. Trans. R. Soc. B* **362**, 1887-1925. (doi: http://dx.doi.org/10.1098/rstb.2006.1881)

Miller SL, Orgel LE. 1974 *The Origin of Life on the Earth* Englewood-Cliffs: Prentice-Hall

Nakada HI, Weinhouse S. 1953 Non-enzymatic transamination with glyoxylic acids and various amino acids. *J. Biol. Chem.* **204**, 831-836. (doi:)

Nitschke W., Russell MJ. 2013 Beating the acetyl coenzyme A-pathway to the origin of life. *Phil. Trans. R. Soc.* **B 368**, 20120258. (doi: http://dx.doi.org/10.1098/rstb.2012.0258)

Orgel LE (1968) Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381-393.

Orgel LE. (2008) The implausibility of metabolic cycles on the Prebiotic Earth. *PLoS Biol.* **6**(1) e18 (doi: http://dx.doi.org/101371.journal.pbio.0060018)

Petrov AS, Bernier CR, Hsiao C, Norris AM, Kovacs NA, Waterbury CC et al. 2014 Evolution of the ribosome at atomic resolution. *Proc. Nat. Acad. Sci USA* **111**, 10251-10256 (doi: http://dx.doi.org/10.1073/pnas.1407205111/-/DCSupplemental)

Piper MO, Hong S-P, Ball GE, Dawes IW. 2000 Regulation of the balance of one-carbon metabolism in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **275**, 30987-30995.

(doi: http://dx.doi.org/10.1074/jbc.M004248200)

Pitsch S, Wendeborn S, Krishnamurthy R, Holzner A, Minten M 2003 The beta-D-ribopyranosyl-(4'-2')-oligonucleotide system(pyranosyl-RNA): Synthesis and resume of base-pairing properties. *Helv. Chim. Acta* **86**, 4270-4363. (doi: http://dx.doi.org/10.1002/hlca.200390316)

Pizzarello S, Weber AL. 2004 Prebiotic amino acids as asymmetric catalysts. *Science* **303**, 1151. (doi: http://dx.doi.org/10.1126/science.1093057)

Poletto M, Pistor V, Zattera AJ. 2013 Structural characteristics and thermal properties of native cellulose. In *Cellulose – Fundamental Aspects* Eds. Van de Ven T, Godbout L. InTech, Chp. 2, 45-68. http://intechopen.com/books/cellulose-fundamental-aspects. (doi: http://dx.doi.org/10.5772/50452)

Powner MW, Gerland B, Sutherland JD. 2009 Synthesis of activated pyrimidine ribonucleotides in prebiotically conditions. *Nature* **459**, 239-242 (doi: http://dx.doi.org/10.1038/nature08013)

Proskurowiski G, Lilley MD, Seewald JS, Frith-Green GL, Olson EJ, Lupton JE,Sylva SP, Kelley DS. 2008 Abiogenic hydrocarbon production at Lost City hydrothermal field. Science 319, 604-607. (doi: http://dx.doi.org/10.1126/science.1151194)

Research News: Max Planck Institute for Biophysical Chemistry 2011 How to remain flexible: The versatile active site of the ribosome. pp. 1-3. Based on Kuhlenkoetter S, Wintermeyer W, Rodnina M. 2011 Different substrate-dependent transition states in the active site of the ribosome. *Nature* (online) (doi: http://dx.doi.org/10.1038/nature10247)

Robinson AB, Scotchler JW, McKerrow JH. 1973 Rates of nonenzymatic deamination of glutaminyl and asparaginyl residues in pentapeptides. *J. Am. Chem. Soc.* **95**, 8156-8159. (doi: http://dx.doi.org/10.1021/ja00805a032)

Rodnina MV, Beringer M, Wintermeyer W. 2006 Mechanism of peptide formation on the ribosome. *Quart. Rev Biophys.* **39**, 203-225. (doi: http://dx.doi.org/10.1017/S003358350600429X)

Rodnina MV, Beringer M, Wintermeyer W. 2007 How ribosomes make peptide bonds. *Trends Biochem. Sci.* **32**, 20-26 (doi: http://dx.doi.org/10.1016/j.tibs2006.11.007)

Russell MJ, Daniel RM, Hall AJ. 1993 On the emergence of life via catalytic iron-sulphide membranes. *Terra Nova* **5**, 343-347. (doi: http://dx.doi.org/10.1111/j.1365-3121.1993.tb00267.x)

Safont VS, Oliva M, Andres J, Tapia O. 2000 Alternative pathways for the C2-C3 bond cleavage and C2 configuration inversion processes for the Rubisco-catalyzed carboxylation sequence. *Chem. Phys. Lett.* **318**, 361-369. (doi: http://dx.doi.org/10.1016/S0009-2614(00)00050-6)

Saks ME, Sampson JR 1995 Evolution of tRNA recognition systems and tRNA gene sequences. *J. Mol. Evol*. **40**, 509-518.

Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Ngyun KA, Su SJ, Windus TL, Duouis M, Montgomery JA. 1993 Gamess general atomic and molecular electronic structure system. *J. Compu Chem.* **14**, 1347-1363.

(doi: http://dx.doi.org/10.1002/jcc.540141112)

Schmitt E, Moulinier L, Fujiwara S, Imanaka T, Thierry J-C, Moras D. 1998 Crystal structure of aspartyl-tRNA synthetase from *Pyrococcus kodakaraensis*: archaeon specificity and catalytic mechanism of adenylate formation *EMBO J*. **17**, 5227-5127. (doi: http://dx.doi.org/10.1093/emboj/17.17.5227)

Schomburg I. 2012 Cofactors and vitamins. In *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, eds.,Schomburg D, Michal G, 2nd Ed., Hoboken: John Wiley & Sons. pp 133-149.

Schoning K, Scholz P, Gontha S, Wu X, Krishnamurthy R, Eschenmoser A. 2000 Chemical etiology of nucleic acid structure: The alpha-threofuranosyl-(3' 2')oligonucleotide system. Science 290, 1347-1351. (doi: http://dx.doi.org/10.1126/science.290.5495.1347)

Service RF. 2013 The life force. *Science* **342**, 1032-1034 (doi: http://dx.doi.org/10.1126/science.342.6162.1032)

Sgrinani J, Magistrato A 2012 The structural role of Mg^{2+} ions in a class I RNA polymerase ribozyme: a molecular simulation study. *J. Phys. Chem. B* **116**, 2259-2268 (doi: http://dx.doi.org/10.1021/jp206475d

Sheppard K 2007 *RNA-dependent Biosynthesis of Glutamine in Bacteria and Archaea* Thesis, Yale Uni., New Haven

Soai K, Shibata T, Morioka H, Choji K. 1995 Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. Nature 378, 767-768. (doi: http://dx.doi.org/10.1038/378767a0)

Sousa FL, Martin WL. 2014 Biochemical fossils of the transition from geoenergetics to bioenergetics in prokaryote one carbon compound metabolism. Biochim. Biophys. Acta 1837, 964-981. (doi: http://dx.doi.org/10.1016/j.bbabio.2014.02.001)

Sponer J, Mladek A, Sponer J, Svozil D, Zgarbova M, Banas P, Jurecka P, Otyepka M. 2012 The DNA and RNA sugar-phosphate backbone emerges as the key player. An overview of quantum-chemical, structural biology and simulation studies. *Phys. Chem. Chem. Phys.* **14**, 15257-15277. (doi: http://dx.doi.org/10.1039/c2cp41987d)

Stahley MR, Strobel SA. 2005 Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science* **309**, 1587-1590. (doi: http://dx.doi.org/10.1126/science.1114994)

Suenaga M. 2005 Facio: New computational chemistry environment for PC Gamess. *J. Comput. Chem. Japan* **4**, 25-32 (doi: http://dx.doi.org/10.2477/jccj.4.25)

Tamura K, Schimmel PR. 2006 Chiral-selective aminoacylation of an RNA minihelix:Mechanistic features and chiral suppression. *Proc. Natl. Acad. Sci USA* **103**, 13750-13752. (doi: http://dx.doi.org/10.1073/pnas.0606070103)

Tamura M, Holbrook SR. 2002 Sequence and structural conservation in RNA ribose zippers. *J. Mol.*

Biol. **320**, 455-474. (doi: http://dx.doi.org/10.1016/S0022-2836(02)00515-6)

Taylor FJR, Coates D. 1989 The code within codes. *BioSystems* **22**, 177-187. (doi: http://dx.doi.org/10.1016/0303-2647(89)90059-2)

Trobro S, Aqvist J. 2005 Mechanism of peptide bond synthesis on the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12395-12400. (doi: http://dx.doi.org/10.1073/pnas.0504043102)

Ulyanov NB, James TL. 2010 RNA structural motifs that entail hydrogen bonds involving sugar-phosphate backbone atoms of RNA *New J Chem* **34**, 910-917. (doi: http://dx.doi.org/10.1039/b9nj00754g)

Wächtershäuser G. 1992 Groundworks for an evolutionary biology; the iron-sulfur world. *Prog. Biophys. Mol. Biol.* **58**, 85-201 (doi: http://dx.doi.org/10.1016/0079-6107(92)90022-X)

Ward PD, Brownlee D. 2000 *Rare Earth: Why Complex Life is Uncommon in the Universe* New York: Copernicus

Walker JCG. 1985 Carbon dioxide on the early Earth *Orig. Life Evol. Biosph.* **16** 117-127. (doi: http://dx.doi.org/10.1007/BF01809466)

Weber A L. 2001 The sugar model: catalysis by amines and amino acid products. *Orig. Life Evol. Biosph.* **31**, 71-86. (doi: http://dx.doi.org/10.1023/A:1006750423942)

Weber AL. 2007 The sugar model: autocatalytic activity of the triose-ammonia reaction. *Orig. Life Evol.Biosph.* **37**, 105-111. (doi: http://dx.doi.org/10.1007/s11084-006-9059-9)

White HB. 1972 Coenzymes as fossils of an earlier metabolic state. *J Mol. Evol.* **7**, 101-104. (doi: http://dx.doi.org/10.1007/BF01732468)

Wickler W. 1968 *Mimicry in Plants and Animals* New York: McGraw Hill

Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. 2004 Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281-286. (doi: http://dx.doi.org/10.1038/nature02362)

Woese CR 1987 Bacterial evolution. *Microbiol. Rev.* **51**, 221-271.

Wünschiers R. 2012 Nucleotides and nucleosides. In *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, eds., Michal G, Schomburg D, 2nd Ed., Hoboken: John Wiley & Sons. pp 124-133.

Yakhnin AV. 2013 A model for the origin of life through rearrangements among prebiotic phosphodiester polymers. *Orig. Life Evol. Biosph.* **43**, 39-47. (doi: http://dx.doi.org/10.1007/s11084-012-9321-2)

Zhang J, Peritz A, Meggers E. 2005 A simple glycol nucleic acid. *J. Am. Chem. Soc.* **127**, 4174-4175. (doi: http://dx.doi.org/10.1021/ja042564z)

Zurek J., Bowman AL, Skolaski WA, Mulholland AJ 2004 MM and QM/MM modeling of threonyl-tRNA synthetase. *Struc. Chem.* **15**, 405-414 (doi: http://dx.doi.org/10.1023/B:STUC.0000037896.80027.2c)